Diversity patterns of honey bee (*Apis mellifera* L.) populations from the archipelago of the Azores: insights from mtDNA and wing geometric morphometrics

Helena Mendes Ferreira
Mestrado em Biodiversidade, Genética e Evolução
Departamento de Biologia
2017

Orientador
Maria Alice Pinto, Professor Doctor, Polytechnic Institute of Bragança (IPB), Mountain Research Centre (CIMO) - Bragança, Portugal

Co-orientador
Miguel Carneiro, Doctor, Research Centre in Biodiversity and Genetic Resources (CIBIO/ InBIO) - Vairão, Portugal

Centro de Investigação de Montanha

CIBIO
Centro de Investigação em Biodiversidade e Recursos Genéticos

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

“We are all in the gutter, but some of us are looking at the stars.”

Oscar Wilde

Acknowledgements

Firstly, I would like to thank to my thesis supervisor professor Maria Alice Pinto for accepting me and for the extraordinary guidance through this study. I am deeply grateful for everything you have done for me, mainly for always being available to clarify all my doubts, for being patient and for everything I had the opportunity to learn with you. I had the opportunity to grow as a person and, above all, to grow as a professional. It was an honor to have you as my supervisor.

Secondly, I would also want to deeply thank Miguel Carneiro, for having accepted to be my co-supervisor and for all the good advices during my thesis writing, mainly during the final phase.

I would like to express my feeling of deep gratitude to Professor Tiago Francoy for having accepted the challenge to teaching me geometric morphometrics, and for the extraordinary guidance through all this process.

I also want to express my deepest gratitude to all veterinarians and technicians of the "Direção Regional da Agricultura e Desenvolvimento Rural dos Açores" who have played a key role in this study by collecting all the samples across the Azores , namely: Franck Aguiar, Nuno Salvador, Janyne Sousa, Ivan castro, Célia Mesquita, Ana Jorge, José Dias, Paulo, Rico, Pedro Leal, Vagner Paulos, Luis Xavier, Luís Silva, Martins Silva, Carlos Gouveia, Ana Carina Coimbra, João Ramos, João Arruda, Edgardo Melo, João Luís, and Moniz da Ponte. Special thanks are due to Paula Vieira for the coordination of the sampling and for the unconditional support. Finally, I would like to thank the "Director Regional de Agricultura", Dr. Fernando Sousa, for authorizing the sampling.

I am so grateful to my friends and also my work colleagues: Dora, Cátia, Julio, Andreia, Paulo (the best beekeeper ever), Fernando, Angelo and the newcomer Miquelina. Thank you for always pushing me and motivating me and for all the support, love and beers (right Andreia?) in the bad and also in the good moments. You were incredible! I could not ask for better friends, with whom I took great pleasure in working and sharing ideas and knowledge. Dora and Cátia, thank you for being always by my side during this long journey, and for the super great time we spent together in that yellow attic. Dora, without you I would never have known the fantastic world of honey bees. Thank you for never giving up on me, and always believe in me and my capacities.

A thank you to all the people with whom I had the opportunity to meet during the time I was in Brazil and especially for the warm welcome. I wanted to express my gratitude to all researchers and technicians in LabApis in Ribeirão Preto, especially

those who helped me and contributed in some way to my work: Clau, Jaque Elisa, Rô Jairo and Ju and Vanessa e Joyce. Clyce, thank you for accepting me with open arms, for patiently having clarified all my doubts and sharing all your knowledge with me. You were certainly one of the most important people during this journey. Thank you for all the love and support.

I would also like to show my gratitude to my lovely friends, Ilda, Marta and Sara, for having been patient with me, and always having the right words of support and motivation which contributed a lot for this work. Ilda, thank you so much for always being available to me when I most needed.

I would like to thank to my friend Rui, for all the scientific debates and the thousand of discussions about the global problems. You know that I always have the best arguments and you never admit, but you know that I'm always right. Thank you for always listening all my problems and concerns.

A special thank to my parents, to whom I owe the privilege of being able to do this master's and this thesis. Thank you for always believing in me and for encouraging me to pursue my dreams.

A huge and special thank to my sisters, Joana and Carla, my brother Ricardo and to Lúcia and Fred. Thank you for always being available to me, for all the love you shared. Thank you for tolerating my bad mood whenever things did not go as expected. Joana, I'm so sorry for all the invitations I had to refuse to go to the cinema, to go shopping, to go for a walk, to go for a coffee. I promise to reward you.

Thanks to my lovely sister Ana. You are a great part of me, and it is in you that I always seek the strength to continue my journey. You cannot imagine how grateful I am to have had the opportunity to meet you and be a sister. Thank you for always light all my steps. I miss you so much.

A huge thank to my little nephews, Maria and João. Thank you for always distracting me from my work, and reminding me of how good it was to be your age. I loved sharing this journey of my life with you. You were undoubtedly the best breaks of my work.

I would like to thank to my sweet grandparents, for the love, care and support.

Finally, I want to thank to one of the most special person I know: my boyfriend Raul. It was great pleasure to make this journey by your side. Thank you for selfless love and for support when I thought it was impossible to move on. No word I can pronounce will express my feeling of gratitude to you. This was just the beginning of our lives.

This study was partially funded by the research project "BEEHEAL: Promoting bee health for sustainable agriculture". This project is funded through the ARIMNet2

(2016) Call by the following funding agencies: INIA (Spain), ANR (France), MOARD (Israel), and FCT (Portugal). ARIMNet2 (ERA-NET) has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 618127.

Resumo

Os padrões de diversidade genética de populações de abelha mellifera (*Apis mellifera L.*) são praticamente desconhecidos nos Açores, contrariamente ao de outros arquipélagos da Macaronésia, como as ilhas Canárias e a Madeira. De forma a tentar preencher esta lacuna, um total de 473 colónias de abelha melífera foram amostradas ao longo das oito ilhas dos Açores ocupadas por esta mesma abelha para posterior análise mitocondrial (mtDNA) e variação morfométrica das asas. A origem do mtDNA foi obtida usando-se a técnica PCR-RFLP da região intergénica $tRNA^{leu}$-cox2, técnica conhecida como teste *Dra*I. As diferenças morfológicas na forma das asas das populações das diferentes ilhas foram analisadas através da morfometria geométrica, usando-se 19 pontos anatómicos distribuídos ao longo da venação das asas. Para além das 473 colónias dos Açores, foram avaliados os padrões morfométricos de uma coleção de referência contendo 711 colónias de *A. m. iberiensis*, 11 colónias de *A. m. ligustica*, e 15 colónias de *A. m. carnica.*

Os padrões maternos foram razoavelmente congruentes com os padrões morfológicos obtidos para as asas. O teste *Dra*I permitiu a identificação de 13 haplótipos diferentes, incluindo um novo haplótipo (A64'). Dos 13 haplótipos, 10 pertencem á linhagem Africana (linhagem A), dois à linhagem da Europa ocidental (linhagem M) e um à linhagem da Europa do leste (linhagem C). Entre as sub-linhagens Africanas conhecidas, a sub-linhagem A_{III} foi a mais frequente (62.38%) e apenas 3% da sub-linhagem A_{I} foi encontrada no Faial. Os dados do mtADN sugerem introduções históricas de colónias de abelha melífera nos Açores, originárias do norte de Portugal, onde uma frequência elevada de haplótipos de ascendência da sub-linhagem A_{III} é encontrada. Para além das introduções históricas, os padrões maternos das populações Açoreanas parecem ter sido moldados por recentes importações de raças comerciais de linhagem C, como é sugerido pelas elevadas frequências de haplótipos de linhagem C encontrados no Pico, Faial e Graciosa.

Os resultados obtidos para a morfometria geométrica revelaram o poder da venação das asas para diferenciar as populações nos Açores e também permitiram distingui-las das subespécies da coleção de referência. Os padrões morfométricos das asas mostraram que, no cômputo geral, as populações dos Açores exibem uma relação próxima com *A. m. iberiensis,* contudo, algumas populações, especialmente da Graciosa, tendem a ficar agrupadas com *A. m. ligustica* e *A. m. carnica*.

O efeito fundador resultante das introduções de *A. mellifera* em períodos passados, o ambiente insular particular, a barreira ao fluxo de genes devido ao isolamento geográfico e o fluxo genético associado à atividade apícola mais recente

por parte do ser humano, são fatores não mutuamente exclusivos que possivelmente moldaram os padrões de diversidade observados atualmente nos Açores.

Palavras-chave: *Apis mellifera*, abelha melífera, Arquipélagos dos Açores, mtADN, teste *Dra*I, morfometria geométrica das asas, introgressão

Abstract

The patterns of genetic diversity of the honey bee (*Apis mellifera L.*) populations from the Azores are virtually unknown, contrary to other Macaronesian archipelagos such as the Canary Islands and Madeira. In an attempt to fill this gap, a total of 473 honey bee colonies were sampled across the eight islands of the Azores occupied by honey bees, which were then surveyed for mitochondrial DNA (mtDNA) and wing morphometric variation. The mtDNA origin was assessed using a PCR-RFLP technique of the intergenic $tRNA^{leu}$-cox2 region, known as the *Dra*I test. Morphological differences in the wing shape among the populations of the different islands were analyzed with the geometric morphometric approach using 19 anatomical landmarks distributed along the wings venation. In addition to the 473 colonies of the Azores, morphometric patterns were assessed on a reference collection comprising 711 colonies of *A. m. iberiensis*, 11 colonies of *A. m. ligustica*, and 15 colonies of *A. m. carnica*.

Maternal patterns were reasonably congruent with the morphological patterns obtained with the wings. The *Dra*I test allowed identification of 13 different haplotypes, including a novel one (A64'). Of the 13 haplotypes, ten belong to the African lineage (lineage A), two to the western European lineage (lineage M), and one to the eastern European lineage (lineage C). Among the known African sub-lineages, sub-lineage A_{III} was the most frequent (62.38 %) and only 3% of sub-lineage A_I was found in Faial. The mtDNA data suggest that historical introductions of honey bee colonies in the Azores originated from the northern part of continental Portugal, where a high frequency of haplotypes of sub-lineage A_{III} ancestry is found. In addition to historical introductions, the maternal patterns of the Azorean populations have seemingly been shaped by recent importations of commercial breeds of C-lineage ancestry, as suggested by the high frequency of C-lineage haplotypes found in Pico, Faial and Graciosa.

The results obtained with geometric morphometrics revealed the power of the wing venation to discriminate different populations in the Azores, and also to distinguish them from the subspecies of the reference collection. The wing morphometric patterns showed that, overall, populations from the Azores exhibited a closer relationship with *A. m. iberiensis*; however some populations, especially those from Graciosa, tend to be clustered closer to *A. m. ligustica* and *A. m. carnica*.

The founder effect resulting from introductions of *A. mellifera* in historical times, the particular insular environment, the barrier to gene flow due to geographical isolation, on one hand, and the contemporary human-assisted gene flow associated with beekeeping activity, on the other hand, are non-mutually exclusive factors that have possibly shaped the diversity patterns observed today in the Azores.

Key-words: *Apis mellifera*, honey bees, Azores Archipelago, mtDNA, *Dra*I test, wing geometric morphometrics, introgression

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

AT: Atlantic Transect

bp: Base pairs

CIMO: Centro de Investigação de Montanha
COX1 gene or ***COI gene***: Cytochrome Oxidase subunit I
COX2 gene or ***COII gene***: Cytochrome Oxidase subunit II
CT: Central Transect
CVA: Canonical Variate Analysis
CV1: Canonical Variate1
CV2: Canonical Variate 2

DFA: Discriminant Function Analysis
DNA: Deoxyribonucleic Acid
DraI: Restriction enzyme from *Deinococcus radiophilus* that recognizes TTT^AAA sites
DWV: Deformed Wing Virus

E2: Reverse primer

GenBank: Genetic sequence database

H2: Forward primer

Iberian honey bee: *Apis mellifera iberiensis*

km: Kilometer
km2: Square kilometer

Lineage A: *African lineage*
Lineage C: *Eastern Europe lineage*
Lineage M: *Western and Northern European lineage*
Lineage O: *Near East and Asia lineage*

MT: Mediterranean Transect

mtDNA: mitochondrial DNA

Mya: Million years ago

PC1: Principal Component 1

PC2: Principal Component 2

PCA: Principal Component Analysis

PCR: Polymerase Chain Reaction

PCR - RFLP: Polymerase Chain Reaction - Restriction Fragment Length Polymorphism

P-value: It is a statistical probability of Karl Pearson

RFLP: Restriction Fragment Length Polymorphism

SNPs: Single Nucleotide Polymorphisms

tRNA^{leu} gene: *transfer RNA leucine gene*

X: Cartesian Coordinate X

Y: Cartesian Coordinate Y

': Haplotype with three Q elements

%: Percentage

3′ UTR: 3′ Untranslated Region

5′ UTR: 5′ Untranslated Region

1. Introduction

1.1 Brief introduction to the genus Apis

Insects are the most successful and diverse organisms in the world, providing vital ecosystem services as pollination, pest control, decomposition, among others (Losey et al. 2006). Among the insect pollinators, the highlight goes to the bees, which are responsible for pollination of both agricultural crops and wild plants (Potts et al. 2010). The Western honey bee, *Apis mellifera* Linnaeus 1758, is one of the main pollinators with a high ecological and environmental impact, besides the great economic importance, and is therefore spread across the entire world (Klein et al. 2007; Potts et al. 2010; Breeze et al. 2011; Han et al. 2012). While honey bees and humans present currently a strict relationship, the oldest records date back to 7000 years ago in the cave paintings in Spain that already demonstrate humans collecting honey (Crane 1999).

The honey bee belongs to the genus Apis, which includes ten species divided into three groups: the cavity-nesting bees (*A. mellifera*, *A. cerana*, *A. koschevnikov*, A. nuluensis, *A. nigrocincta*), the giant bees (*A. dorsata*, *A. laboriosa*, *A. binghami*), and the dwarf bees (*A. florae, A. andreniformis*) (Arias and Sheppard 2005; Raffiudin and Crozier, 2007) (Figure 1). Based on phylogenetic analysis, *A. breviligula and A. indica* should be considered as species instead of subspecies and added to the genus (Lo et al. 2010) (Figure 1).

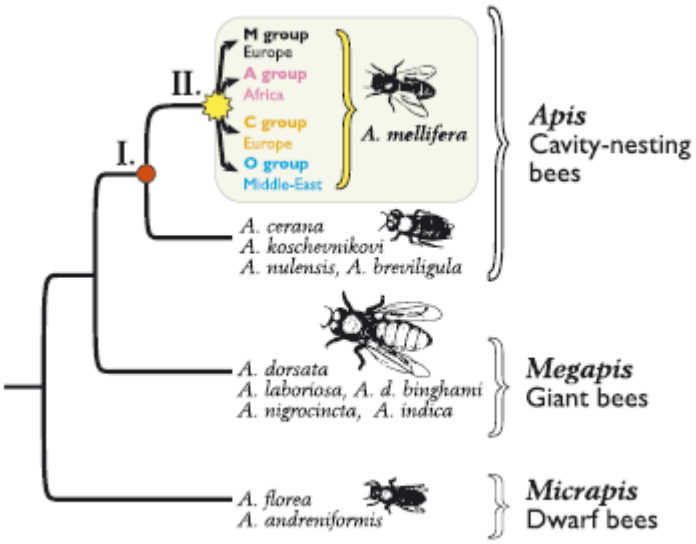

Figure 1. Phylogeny representing the three clades of Apis. All of the 10 extant Apis species apart from *A. mellifera* are found only in Asia. Node I represents the split between *A. mellifera* and other cavity-nesting bees. Node II represents the most recent common ancestor of extant subspecies of *A. mellifera* (Adapted from Han et al. 2012).

Two cavity-nesting bees, *A. mellifera* and *A. cerana*, were considered the most recently derived sister taxa. Based on sequence data from mtDNA and nuclear loci it is estimated that *A. mellifera* diverged from the sibling species *A. cerana* between 6 and 9 million years ago (Cornuet and Garnery 1991; Arias and Sheppard 2005). Apart from A. mellifera, all of the bee species from genus Apis are currently confined to Asia.

Eusocial bees are characterized by the presence of three different castes in the same colony: drones, workers and the queen (Boomsma et al. 2005; Barchuk et al. 2007). Hymenoptera males of the genus Apis are haploid; emerging from unfertilized eggs and die right after mating (Baer 2005; Boomsma et al. 2005). Workers are originated from diploid eggs and are usually sterile. Each colony of A. mellifera possesses normally only one active queen which is able to generate offspring (Barchuk et al. 2007). The nuptial flight of honey bee queens is performed few days after the emergence of the queen and mating takes place in male congregation areas (Baer 2005). Species of the genus Apis are extremely polyandrous. In queens from A. mellifera the number of matting can vary between subspecies (Kraus et al. 2004).

1.2. Evolutionary history of Apis mellifera

After the split from *A. cerana*, *A. mellifera* eventually diversified into numerous subspecies and attained its present distributional range in Europe, Africa, Western and Central Asia (Ruttner 1988). Although mtDNA data suggests that the four evolutionary lineages diverged around 0.7 to 1.3 million years ago (Garnery et al. 1992; Arias and Sheppard 1996), recent estimations based on the genetic analysis of 8.3 million SNPs resulting from whole genome sequencing indicate a split from each other around 300 000 years ago (Walberg et al. 2014).

Initially, 25 subspecies were described by Ruttner (1988) based on morphological traits. The advent of new molecular techniques increased this number to 30 (Ruttner 1988; Hepburn and Radloff 1988; Engel 1999; Sheppard et al. 1997; Sheppard and Meixner 2003; Meixner et al. 2011). The 30 subspecies of *A. mellifera* were grouped in four main evolutionary lineages, showing a particular geographic distribution: African lineage A, Western and northern Europe lineage M, Eastern Europe lineage C and Middle East and Western Asia lineage O (Figure 2). The exactly geographic range and distribution of the different subspecies of *A. mellifera* are showed in the supplementary material (Table S1). Recently a new subspecies of A. mellifera winter tolerant was described in Xinyuan prefecture, located in Xinjiang Uygur Autonomous Region of China and named *A. m. sinisxinyuan (*Chen et al. 2016).

According to morphological and molecular data this new subspecies was placed in the lineage M and the estimated time of divergence between *A. m. mellifer*a and *A. m. sinisxinyuan* is about 132 000 years (Chen et al. 2016). The differences in morphology, behavior and biological traits observed across the natural range of the different subspecies of *A. mellifera*, are the result of historical patterns of isolation and posterior adaptation to distinct habitats (De la Rua et al. 2009). Furthermore, the native range occupied by the western honey bee, *Apis mellifera* was extended worldwide to support pollination and for the honey production (Jaffé et al. 2010; Byatt et al. 2015).

Initially, the subspecies described by Ruttner (1988) were grouped, based on morphometric analysis, into the four evolutionary lineages: A, M, C and O. With the development of molecular techniques and new morphological data, some of the subspecies were reassigned to other lineages, such as the case of *A. m. intermissa*, *A. m. sahariensis,* and *A. m. siciliana*, which were originally placed in lineage M (Ruttner et al. 1978) and later relocated to lineage A (Ruttner 1988). Mitochondrial data was generally concordant with the morphological lineages; however, the mtDNA did not support the existence of lineage O, as it does not distinguish between C and O haplotypes (Cornuet and Garnery 1991; Garnery et al. 1992). Indeed, while morphometric analyses placed *A. m. cypria*, *A. m. anatoliaca* and *A. m. caucasia* in the lineage O (Ruttner 1988; Kandemir et al. 2011), mitochondrial data suggested that they belong to lineage C (Garnery et al. 1992; Smith et al. 1997; Kandemir et al. 2006). Therefore, later studies considered only three maternal lineages considered (A, M and C) while morphometry identified four lineages (A, M, C and O.). Later on, a new pattern of mtDNA variation that differed from those previously reported was observed in populations of *A. mellifera* from Near East, and the existence of a fourth maternal lineage was postulated (Franck et al. 2000a; Palmer et al. 2000). This new lineage was named O, which was erroneously associated with the morphological lineage O (Franck et al. 2000a). Recently, with the inclusion of additional geographical data and with new sequence analyses, lineage O was reclassified as an African sub-lineage and renamed Z (Alburaki et al. 2011). Along with the sub-lineage Z, the complexity of African lineage is also reflected by additional three sub-lineages (A_I, A_{II}, A_{III}) described by Franck et al. (2001). Also the maternal lineage Y, identified in populations of *A. m. jemenitica* from northeastern Africa (Franck et al. 2001), is seemingly part of the African (lineage A) maternal variation (Meixner et al. 2013). Analyses based on Single Nucleotides Polymorphism (SNPs) were largely congruent with the existence of the four divergent evolutionary lineages (Whitfield et al. 2006; Han et al. 2012) proposed by morphological data (Ruttner 1988).

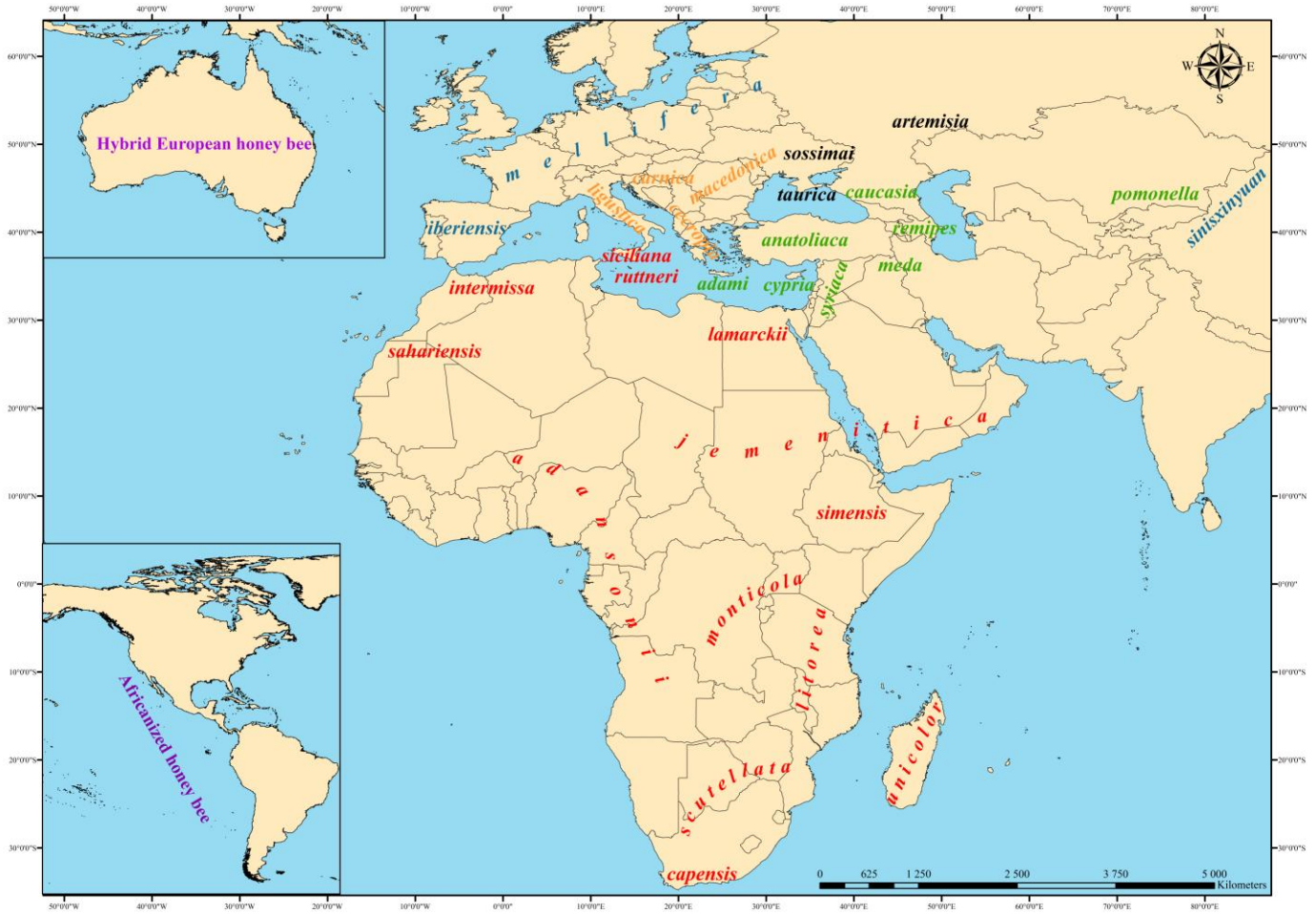

Figure 2. Map with the geographical distribution of the 31 subspecies of *A. mellifera*. The color of the names indicated the respective evolutionary lineage: A-African lineage (red), M-Western and Northern European linage (blue), C- Eastern European lineage (orange) and O- Middle East and Western Asia lineage (green). *A. sossimai* (Engle 1999), *A. m. taurica* (Apaltov 1938) and *A. m. artemisia* (Engle 1999) are subspecies with sparse information on their evolutionary lineage (black) (Adapted from Chávez-Galarza et al. 2016a).

Regarding the origin of *A. mellifera*, there are some doubts that several studies have not been yet able to clarify and so far there are multiple scenarios under discussion. The first scenario of the origin of the evolutionary lineages was proposed by Ruttner (1978) that using morphological characters stated that the origin of the lineages was in the Middle East or Northeast Africa, from where the Europe was later colonized through two routes: via Western and Northwestern Africa to the Iberian Peninsula and via Middle East to the Balkans. This theory fits with molecular analyses, namely 1029 SNPs from Whitfield et al. (2006) reanalyzed by Han et al. (2012). Using different genetic parameters conclude that the origin of *A. mellifera* occurred probably in Asia, close where other Apis species are currently found (Han et al. 2012) (Figure 3A).

Also Garnery et al. (1992), based on mtDNA analyses proposes a Middle Eastern origin, but unlike Ruttner (1978) does not propose a colonization of Europe via western route and migration across the strait of Gibraltar, once the phylogenetic tree groups A and C lineage rather than M. The recent discovery of the new subspecie *A. m. sinisxinyuan* in the Tian Shan Mountains raises questions about the origin of the M lineage, suggesting a colonization of Europe through Asia, supporting the scenario proposed by Garnery et al. (1992) (Figure 3B).

Studies conducted by Whitfield et al. (2006) with basis on 1136 SNPs derived from the nuclear genoma support an origin of the honey bee in Africa however with the lineage A giving origin to the M lineage, who expanded later through the Western. According Whitfield et al. (2006), lineage C colonized the Europe while O lineage spread into Asia (Figure 3C).

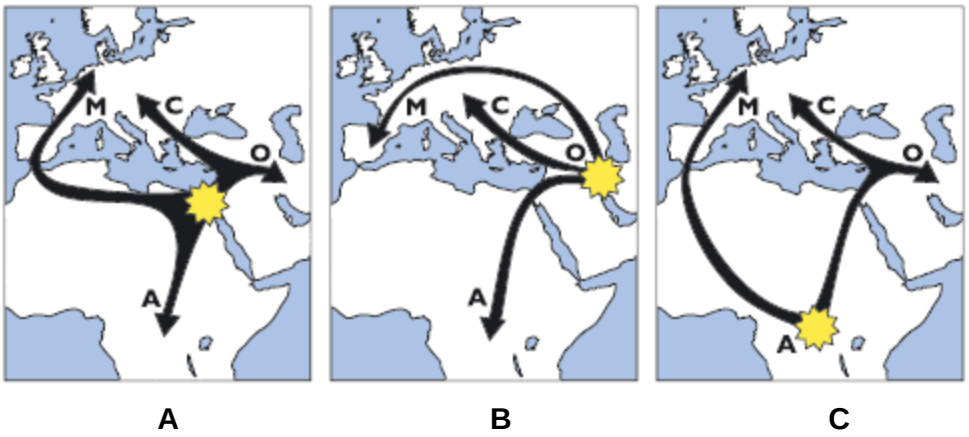

Figure 3. Three of the four hypotheses proposed for the origin of A. mellifera: (A) An expansion from the Middle East, involving colonization of Europe via two routes one eastern and one western was first suggested by Ruttner (1978) on the basis of morphometric analyses. This theory fits with the molecular analyses by Han et al. 2012 (B) An expansion from the Middle East, which did not involve the western colonization route into Europe was suggested on the basis of trees constructed from mtDNA (Garnery et al. 1992). (C) An origin in Africa was proposed by Wilson (1971) and an expansion out of Africa via both an eastern and western route was suggested by the analysis of >1000 SNPs by Whitfield et al. (2006) (Adapted from Han et al. 2012).

In the most recent study Cridland et al. (2017) combined whole genome datasets generated by Harpur et al. (2012) and Walberg et al. (2014) and added other whole genomes to understand the evolutionary history of this subpecies. Results highlight a possible origin of the honey bee in the Middle East or in Northeastern Africa (Cridland et al. 2017), supporting the earlier hypothesis of Ruttner (1988), and also molecular analyses (Han et al. 2012; Walberg et al. 2014). However instead of an ancient split between the A and O lineages, Cridland et al. (2017) proposed that the M, C and O are all originated in the African continent.

1.3. Mitochondrial DNA surveys in the honey bee

The mitochondrial DNA (mtDNA) of the honey bee was completely sequenced in 1992 (Crozier and Crozier, 1993). The mtDNA is a circular molecule ranging from 16500 to 17000 bp in honey bees, according to the observed length variability in different regions of the molecule (Cornuet and Garnery, 1991). Because the maternally inherited, all individuals in the colony (workers and drones) share the mtDNA with their mother queen. Thus, with a study of a single individual is possible to identify the maternal ancestry of individual colonies and also to understand the patterns of gene flow or introgression among hybridizing population (De La Rúa et al. 2009). Moreover, the non-recombinant nature of mtDNA makes data interpretation straightforward.

A variety of markers, including RFLP (Restriction Fragment Length Polymorphism), PCR-RFLP (Polymerase Chain Reaction- Restriction Fragment Length Polymorphism) and direct sequencing has been used to assess the mtDNA variation within honey bee (Meixner et al.2013). In the RFLP technique, the digestion with restriction endonucleases has been performed in the entire mitochondrial genome. Although a battery of restriction enzymes (*Hinf*I, *Acc*I, *Ava*I, *Bcl*I, *Bgl*II, *Eco*RI, *Hinc*II, *Hind*II, *Hind*III, *Nde*I, *Pst*I, *Pvu*II, *Xba*I) has been applied in to studies of A. mellifera, and roughly differentiated between the three maternal lineages A, M and C, they do not prove to be diagnostic markers to indentify the different subspecies (Meixner et al. 2013). Later, a different technique was developed, consisting of RFLPs of PCR fragments, requiring fewer amounts of DNA and allowing the amplification of specific regions. Several mitochondrial genes in different subspecies were digested with a distinct range of restriction endonucleases (see supplementary material Table S2). The intergenic $tRNA^{leu}$-cox2 region, described by Cornuet and Garnery (1991) is the most used mitochondrial region. Since the studies developed by Garnery et al. (1993), the *Dra*I test, which consists in the PCR amplification of the intergenic $tRNA^{leu}$-cox2 region

followed by the digestion with the *Dra*I, becomes widely used to assess the genetic diversity within honey bee populations, associated with the relatively low cost and scoring simplicity (Reviewed by Meixner et al. 2013).

1.3.1 *Dra*I test and the intergenic tRNA^{leu}-cox2 region

The intergenic tRNA^{leu}-cox2 region has been assessed mostly by using the popular *Dra*I test, a PCR-RFLP assay which consists of PCR amplification of the intergenic fragment followed by the digestion with the *Dra*I restriction endonuclease, as described by Garnery et al. (1993). The intergenic region is a non-coding region with two distinct nucleotide sequences known as the P and Q elements, which is determined by the number of Q repeats and the forms of the P elements. The P element ranges from ~53 to 68 bp and exhibits three forms: P_0, P, P_1. The ~15 bp difference between P_0 to P_1 is due to a deletion in the 3´end of P element. The Q element, which varies between ~194 and 196 bp, can be repeated in tandem one to five times, although these repeats are not lineage specific (Garnery et al. 1993; De La Rúa et al. 1998; Franck et al. 1998, 2001; Alburaki et al. 2011; Rortais et al. 2011).

The intergenic region tRNA^{leu}-cox2 has been assessed over the years in the study of honey bees providing a powerful and informative content, which has allowed discrimination of the three main evolutionary lineages (Garnery et al. 1993; Franck et al. 1998). Moreover, the DraI test was able to separate the four different sub-lineages known within the African lineage: A_I; A_{II}; A_{III} and Z (Franck et al. 2001; Alburaki et al. 2011). Lineage A presents two forms of the P element: P_0 is characteristic of the sub-lineage A_I, A_{II} and Z, while P_1 is observed in sub-lineage A_{III}. The element P_0 differs from P_1 by a deletion of 15 bp in P_1 in the P form at the 3' end of the P element. Further distinction between sub-lineages A_I, A_{II} and Z is possible due to the number of *Dra*I recognition sites. Sub-lineage A_{II} is differentiated from sub-lineage A_I by the absence of a *Dra*I restriction site in the 5' end of the first Q element, while sub-lineage Z exhibits an additional restriction site in the middle of the first Q element. Lineage C is characterized by the absence of the P element and only presents one Q element. Lineage M carries the P element, which has a large deletion of 13 bp in its middle, besides the two *Dra*I restriction sites in the first Q element, as observed in sub-lineage Z. Although the *Dra*I test exhibit high levels of polymorphism that allowed the identification of numerous haplotypes from M and C lineages and African sub-lineages A_I, A_{II}, and A_{III}, it is unable to distinguish the honey bees at the subspecie level, since some haplotypes are carried by different subspecies (e.g. A1 haplotype was observed

in *A. m. iberiensis* and *A. m. adansonii*; C1 was observed in *A. m. carnica* and *A. m. ligustica*).

1.3.2. *MtDNA* surveys in the Iberian honey bee *A. m. iberiensis*

The Iberian Peninsula is considered one of the most important glacial refugia in Europe during the Quaternary, offering a suitable habitat for many taxa (Goméz and Hunt, 2007). Events associated with this period strongly affected the distribution of species throughout the Iberia Peninsula and has deeply shaped the phylogeographic patterns of many species (Goméz and Hunt 2007). Since the 1990's, many studies have been developed in Iberian Peninsula with the native honey bee (*A. m. iberiensis*) to uncover patterns of genetic diversity. The mtDNA of A. *m. iberiensis* has been extensively surveyed mainly using the *Dra*I test, but also sequence data, reveling a further complexity of honey bee populations from Iberian Peninsula (Franck et al. 1998; Garnery et al. 1998; De La Rúa et al. 2002; Miguel et al. 2007; Cánovas et al. 2008; Pinto et al. 2012, 2013). MtDNA data revealed the occurrence of two divergent evolutionary lineages forming a sharp cline, with lineage M existing mainly in the northeastern half and lineage A in the southwestern half of Iberia (Garnery et al. 1993; Garnery et al. 1998; De La Rúa et al. 2002; Miguel et al. 2007; Cánovas et al. 2008; Chávez-Galarza et al. 2017). The presence of a large diversity of haplotypes from M and A lineage ancestry makes the Iberian Peninsula one of the regions with the highest maternal diversity of in Europe (Cánovas et al. 2008; Pinto et al. 2013; Chavez-Galarza et al. 2017). Other markers have been used in addition to mtDNA including morphology (Cornuet et al. 1989; Arias et al. 2006; Miguel et al. 2010), allozymes (Smith and Glenn 1995; Arias et al. 2006), microsatellites (Franck et al. 1998; Garnery et al. 1998; De la Rúa et al. 2002, 2003; Miguel et al. 2007; Miguel et al. 2010; Cánovas et al. 2011) and SNPs (Chávez-Galarza et al. 2013, 2015). Studies using SNPs reinforced the idea that the Iberian Peninsula may have acted as glacial refuge during glaciations (Chávez-Galarza et al. 2015), as previously speculated by other authors using different molecular markers (Smith et al. 1991; *Franck* et al. 1998; De La Rúa et al. 2002; Miguel et al. 2007; Cánovas et al. 2008).

In addition to the northeastern-southwestern cline, the *Dra*I test also revealed that Atlantic side of the Iberian Peninsula is a hotspot of African maternal diversity (Pinto et al. 2012; 2013). The maternal composition of honey bee populations in Portugal also exhibits a north-south partitioning with northern populations mainly bearing haplotypes of sub-lineage A_{III} ancestry contrasting with the southern populations mostly composed of sub-lineage A_I haplotypes (Pinto et al. 2012; 2013).

Moreover, sub-lineage A_{III} is virtually absent in the African continent (Franck et al. 2001; Pinto et al. 2013) but commonly observed in the Macaronesia Islands, except in Cabo Verde (De La Rúa et al. 1998, 2001a, 2006; Muñoz et al. 2013). Contrasting with sub-lineages A_{III} and A_{II}, the occurrence of sub-lineage A_{II} is less frequent in honey bee population in Iberian Peninsula (Cánovas et al. 2008; Pinto et al. 2013).

1.3.3. *MtDNA* surveys in Macaronesian honey bee

Darwin has turned out the attention to the study of islands since he developed the theory of evolution by natural selection in the finches of the Galápagos Islands (Emerson et al. 2002). The scientific interest of islands is mainly triggered by the fact that they have a small size with oceanic boundaries, displaying a diversity of habitats with high levels of endemism providing natural laboratories for evolutionary studies (Emerson et al. 2002; Losos et al. 2009). Thus, the evolutionary processes underlying genetic diversity of insular populations may be distinct from those in the mainland and the reduced gene flow due to the oceanic barriers and also the genetic drift associated to founder events may lead to the formation of new species (Barton 1996; Emerson et al. 2002). In honey bees, the differences observed in insular populations allowed the characterization of new subspecies as is the case of the *A. m. siciliana* from Sicily (Sinacori et al. 1998), *A. m. adami* from Crete (Bouga et al. 2005), *A. m. ruttneri* from Malta (Sheppard et al. 1997), A. m. cypria from Cyprus (Bouga et al. 2005) and A. m. unicolor from Madagascar (Hepburn and Radloff 1988). Examples of adapted insular populations are described in other Apidae, such as the case of the two endemic Macaronesian bumblebees *Bombus canariensis*, from the Canaries, and *Bombus maderensis*, from Madeira (Erlandsson 1979). The geographic isolation also plays an important role in the development of particular traits in honey bees from islands that made possible their survival and local adaptation. For example, Sinacori et al. (1998) reported that *A. m. siciliana* can reduce or interrupt the brood rearing during summer and also has the ability to control the infestation of the mite *Varroa destructor*. In 1950 Eckert reported that honey bees from the Molokai Island in Hawaii were resistant to American Foulbrood, which is caused by *Paenicillus larvae*, a highly destructive brood disease (Szalanski et al. 2015).

La Rúa et al. (2006) assessed the genetic structure of honey bee populations from the Macaronesian islands of Madeira and São Miguel using the *Dra*I test and microsatellites. MtDNA analysis of the populations from São Miguel revealed the existence of four African haplotypes belonging to sub-lineages A_{III} and A_{II} and confirmed the presence of colonies of C-lineage ancestry. Later in a temporal analysis

of the mitochondrial diversity also assessed by the *Dra*I test, Muñoz et al. (2013) obtained similar results for São Miguel, however with a subtle increase of haplotypes of C-lineage and sub-lineage A_{II} ancestry concurrently with a decrease of sub-lineage A_{III} haplotypes. Yet, these studies on São Miguel certainly are not representative of all the maternal genetic diversity existing across the entire archipelago.

Contrary to the Azores, in the last two decades the Canary Islands have been targeted by extensive sampling and many studies to unveil the genetic diversity of local populations (De La Rúa et al. 1998, 2001, 2002, 2003, Muñoz et al. 2013; Miguel et al. 2016). Not all Canary Islands can support honey bee colonies due to the lack of abundant nectar flows, such as in Lanzarote and Fuerteventura (De La Rúa et al. 1998). The remaining islands (Tenerife, El Hierro; La Gomera, Gran Canaria and La Palma) carry haplotypes mainly belonging to the African sub-lineages A_{III}, in addition to the considerable amounts of lineage C (De La Rúa et al. 1998). For this reason, an African origin was initially suggested for honey bees from Canarias (De La Rúa et al. 1998; 2001), conjecture that proved to be wrong when haplotypes of sub-lineage A_{III} ancestry were found to be abundant in the populations of northern Portugal, supporting instead a Portuguese origin (Miguel et al. 2016). Lineage C was not detected in La Palma in 1998 (De La Rúa et al. 1998), but recent studies report hybridization in middle and western populations, however in low levels (Miguel et al. 2016). In general, the lineages representative of the archipelagos of Canary, Madeira and Azores are of African ancestry, being sub-lineage A_{III} predominant (De La Rúa et al. 1998, 2001, 2002, 2003; Muñoz et al. 2013; Miguel et al. 2016), except in Cape Verde (Franck et al. 2001). Interestingly, this African archipelago only carries haplotypes belonging to sub-lineage AI, similarity to sub-Saharan Africa (Franck et al. 2001).

1.3.4. Signs of hybridization in insular populations of Apis mellifera: consequences and conservation plans

Contrary to other domestic species, the mating in honey bee queens is extremely difficult to control, so the hybridization and introgression between local and introduced honey bees is very common (Franck et al.1998; Jensen et al. 2005; De La Rúa et al. 2009). The honey bees that have been globally imported are derived from the subspecies *A. m. ligustica* and *A. m. carnica*, both belonging to C lineage. These subspecies typically possess a docile nature and high productivity, which makes them favorite for beekeepers around the world (Ruttner 1988, Moritz et al. 2005). However, the introduction of foreign subspecies exposes the native subspecies to introgressive

hybridization, which modifies the gene pool of the local adapted populations leading to the loss of the genetic identity and threatening the survival of local populations across the world (e.g. Pinto et al. 2004; Jensen et al. 2005; De La Rúa et al. 2009; Pinto et al. 2014; Ivanova et al. 2007; Muñoz et al. 2014b). Native subspecies are important reservoirs of local adaptation (De La Rúa et al. 2009) due to the considerable environmental variation, to the weather conditions and flowering seasons but also the occurrence and variation of parasites and pathogens (Meixner et al. 2015). Furthermore, native honey bee exhibits particular behavior and morphological traits that reflect the adaptation to the natural habitats (De La Rúa et al. 2009).

Despite being away from the continental masses, honey bees that inhabit the islands do not become immune to the introduction of foreign subspecies; in fact, isolation makes them even more vulnerable. Populations in the islands are more sensible to the extinction than those in the mainland due to several factors, including human over exploitation, habitat destruction or the problematic introduction of foreign species or diseases (Frankam 2008). Honey bee populations of the Mediterranean Islands are a good example of the harmful effects of the introduction of foreign subspecies. The sicilian honey bee, *A. m. siciliana,* has strongly hybridized with the honey bee from Italy, *A. m. ligustica,* since the beginning of the 20th century, which has almost led to the loss of this subspecies that is well adapted to the harsh environmental conditions of Sicilia (Muñoz et al. 2014b). In the Balearic Islands, honey bee populations from Formentera, Ibiza and Menorca present signs of introgression, which has drastically changed the genetic pool of local populations (De La Rúa et al. 2001, 2009). Although in Canary Islands is observed haplotypes from C lineage (De la Rua et al.1998, 2001, 2002; Muñoz et al. 2013; Miguel et al. 2016), analysis of mitochondrial diversity revealed that São Miguel presents the highest frequencies of C lineage haplotypes, when comparing to the remaining archipelagos of Macaronesia (De La Rúa et al. 2006; Muñoz et al. 2013).

In addition to the loss of the genetic identity, the introduction of foreign queens also exposes the native populations to foreign pathogens (De La Rúa et al. 2009). Until 2000, when were reported massive illegal introductions of foreign subspecies from C lineage in Pico, there were no records of the mite *Varroa destructor*. In the following year, *Varroa* was introduced in Flores, and later, in 2008 arrived at Faial. However, in the remaining islands Varroa is absent which make the beekeeping practice profitable, by not requiring treatment for the mite, which are relatively expensive, and also for the highest commercial value of the products derived from honey bee activity not affect by the mite.

In an attempt to control the introduction of foreign queens, where the local honey bee already exists, local honey bee conservation programs have emerged worldwide to protect local honeybee from genetic introgression. The early studies conducted in Las Palmas did not detect the presence of lineage C (De La Rúa et al. 1998; 2000), which led to the creation in 2001 of a conservation plan named “Honey bee Conservation Project on La Palma”. This also aimed to control the introduction of foreign subspecies, by banning honey bee importations (Muñoz and De La Rúa, 2012; Muñoz et al. 2013; Miguel et al. 2016).

1.4. The importance of the study of morphological traits

The study of shape has played a crucial role in the world of biology as a way of investigating the inherent variation of different organisms. There are several factors that affect the morphology and that can lead to the variability that is observed among the individuals, molded by the strict relation between the organism and the environment (Ricklefs and Miles 1994). Competitive interaction between closely related species, unequal growth and morphogenesis processes, as well as different selective pressures caused by climate or habitat can lead to morphological differentiation (Atchley and Hall 1991; Ricklefs and Miles 1994; Adams and Rohlf 2000; Zelditch et al. 2004). Thus, it is possible to identify different groups of individuals based on morphological traits. Morphometry has therefore been widely used over the years covering a vast spectrum of species contributing to the taxonomic classification of species mostly before the availability of molecular approaches.

1.4.1. Traditional Morphometry

Morphometric studies in honey bees began in 1916 in Russia when Cochlov compared different races of A. *mellifera* from different geographic zones using the length of the tongue (Ruttner 1988). The works of Alpatov (1929) and Goetze (1940) using univariate statistics contributed to describe different subspecies of *A. mellifera*. Later on, DuPraw (DuPraw 1964, 1965) developed multivariate statistical techniques, known as traditional morphometry or multivariate morphometrics. Traditional morphometrics consisted in the application of multivariate statistical analyses to the study of morphological variables as linear distance measurements but also angles and ratios (Adams et al. 2004; Slice 2007). Ruttner (1988) played a crucial role in the identification and classification of *A. mellifera* subspecies based on the concepts of traditional morphometry. The author was responsible for the description of 36 characters, including distances and angles, particularly related with the wing shape

measurements, analyzed with multivariate statistical analysis (Ruttner 1988). It should be noted that studies with traditional morphometry allowed Ruttner (1988) to identify the four evolutionary lineages in *A. mellifera* (A, M, C, and O).

1.4.2. Geometric morphometrics

At the end of the 20th century, a new technique of geometric morphometrics emerged using information on shape variation obtained from homologous landmarks coordinates (Bookstein 1991, Bookstein 1996). Geometric morphometrics requires landmarks describing specific anatomical locations that should be sufficient to cover the shape of the structure, be easily located and identifiable across the individuals (Zelditch et al. 2004; Adam et al. 2013). Homologous landmarks provide a more complete biological interpretation of the results (Rohlf and Marcus 1993). However, sometimes structures possess other landmarks called semi-landmarks that can capture information about the shape in boundary curves, in the end of structures or arbitrary points along an outline (Rohlf and Marcus 1993; Adam et al. 2013). Differences between specimens can be analyzed using several superimposition methods. Procrustes superimposition methods use raw coordinates obtained through homologous landmarks points of the specimens in study in order to extract shape information, which removes variation in size, orientation and position (Goodall 1991; Rolf and Marcus 1993; Bookstein 1996; Zelditch et al. 2004).The aligned landmark configurations generate points that lie in Kendall's shape space in order to capture all possible variations in shape, and then are projected orthogonally into their Euclidean space tangent (Bookstein 1996; Rohlf 1999; Zelditch et al. 2004). This step is essential since conventional tools of multivariate statistical analysis work with the Euclidean linear space.

In contrast to geometric morphometrics, traditional morphometry has several disadvantages mainly by working with linear distance measurements that are usually deeply correlated with size, and many of those distances are not defined by homologous landmarks (Zelditch et al. 2004). Another difficulty is related to the same distances obtained from different shapes and the impossibility of representing them graphically (Bookstein 1994). Moreover, geometric morphometric approach has been replacing the traditional morphometry since it yields a powerful statistical analysis, is more efficient to detect the variations of shape and allows the graphical representation of geometric structures (Bookstein, 1991; Marcus and Corti 1996; Rohlf and Marcus 1993). Geometric morphometric techniques covers several fields in biology, paleontology and systematics and has been applied to a wide number of plants

(Shipunov and Bateman 2005; Viscosi and Cardini 2011; Savriama et al. 2012), invertebrates (Anstey and Pachut 2004; Mutanen and Pretorius 2007; Wappler et al. 2012; Smith et al. 2013) and vertebrates (O'Higgins and Jones 1998; Klingenber et al. 2001; Claude et al. 2004; Clabaut et al. 2007; Foster et al. 2008; Figueiredo et al. 2009; Ginter et al. 2012; Ottoni et al. 2013), including humans (Bookstein at al. 2001; Bruner 2004; Gonzalez et al. 2009; Perez et al. 2006; Harvati 2004; Gomez-Robles et al. 2008).

1.4.2.1. Geometric morphometric in bee wing venation

Morphometric geometric analysis in insects is mainly based upon the characteristics of the wings, such as shape and venation. Results obtained over the years using wing venation provided reliable information and have demonstrated the efficacy of this technique (Francoy et al. 2008, 2009; Matias et al. 2001; Baylac et al. 2003; Villemant et al. 2007; Gumiel et al. 2003; Perrard et al. 2012, 2014; Aytekin et al. 2007; Hoffman and Shirriffs 2002), when compared with traditional morphometry (Tofilsky 2008). Moreover, it was possible to identify new bee species using the wing venation based on the extremely rare bee fossil record (De Meulemeester et al. 2012; Wappler et al. 2012).

Wings are two-dimensional structures, with many anatomical landmarks, formed by homologous intersections of the veins, which can be compared among groups of individuals, populations, subspecies or species in order to appraise genetic diversity (Michez et al. 2009; Baylac et al. 2003; De Meulemeester et al. 2012; Gerard et al. 2015). Because they are flat and rigid structures, wings can be easily handled, which make them useful tools when compared with other organs (Michez et al. 2009; De Meulemeester et al. 2012; Gerard et al. 2015).

In *A. mellifera*, the results obtained through the morphometry allowed to discriminate between subspecies (Tofilski 2008; Kandemir et al. 2011; Barour and Baylac 2016) and also to infer about the effects of the Africanization process in Brazil (Francoy et al. 2008; 2009). In addition, geometric morphometrics of wings is often used along molecular tools such mtDNA, microsatellites and SNPs (Miguel et al. 2010; Oleksa and Tofilski 2015; Miguel et al. 2016) to better understand the evolutionary patterns that shaped variation observed between the different subspecies of *A. mellifera*.

A few studies using geometric morphometrics for the characterization of *A. m. iberiensis* have been performed in the Iberian Peninsula (Miguel et al. 2010; Chávez-Galarza et al. 2016b). In the most recent study results obtained with the variation of

wing shape of workers from 711 colonies collected across the Iberian Peninsula detected a Southwestern-Northeastern cline with the contact of two divergent evolutionary lineages A and M (Galarza et al. 2016b), previously undetected with traditional morphometry (Cornuet and Fresnaye 1989). These results were also largely congruent with by both nuclear (SNPs) and mitochondrial (mtDNA) markers, being able to capture the signature of complex evolutionary processes (Chávez-Galarza et al. 2013, 2015).

Objectives

The present study aims to provide a comprehensive survey of the genetic diversity of the honey bee populations inhabiting the Azores archipelago which will help understanding the relationship between insular populations and *A. m. iberiensis* populations, the putative historical source of the Azorean honey bees, and the impact of contemporary beekeeping activities. The main objectives of this study were:

-To examine the mitochondrial DNA variation of the intergenic $tRNA^{leu}$-cox2 region using the *Dra*I test;

-To characterize the patterns of wing geometric morphometric variability of the honey bee from Azores Archipelago, in comparison with reference samples representing three subspecies, using multivariate statistical analyses;

- To examine how concordant are the patterns revealed by mtDNA and wing geometric morphometrics data;

-To examine the impact of modern beekeeping in the local populations through detection of variation of C-lineage ancestry.

The findings of this study can be used as a baseline to delineate breeding and conservation programs towards a more sustainable beekeeping in the Azores.

The Macaronesian honey bee populations harbor important genetic diversity, yet this diversity has been increasingly threatened by introduced pests and diseases and genetic pollution from imported queens. Conservation of locally adapted diversity is paramount for honey bees to respond to both climate change and invasive pests and diseases.

2. Material and Methods

2.1. Study Area

The Atlantic archipelagos, of volcanic origin, located next to Europe and North Africa are collectively known as Macaronesia. The Azores is the youngest archipelago in the Macaronesia biogeographical region, which also comprises the archipelagos of Madeira, Canaries and Cape Verde (Fernándes-Palacios et al. 2011) (Figure 4). The Azores comprises nine islands and small islets that are located between 37º and 40º N latitude and 25º to 31 º W longitude, laying at 1584 km from the Portuguese mainland, with a total of 2333 km^2. The Azores are located in a volcanically active region, in the junction of the Eurasian and the American plates.

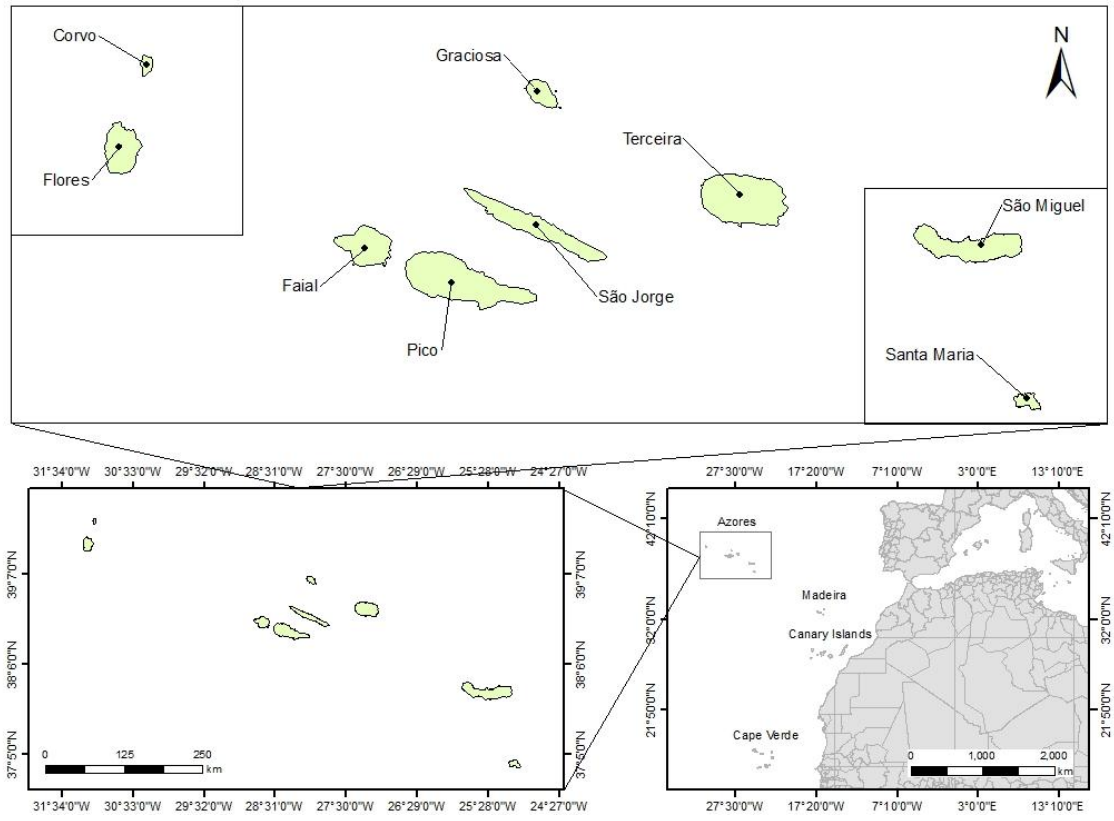

Figure 4. Map of the study area showing the location of the nine islands of the Azores archipelago, located in the Macaronesia biogeographical region.

The Azores is divided into three groups of islands: the Eastern, the Central, and the Western. The Eastern group includes São Miguel and Santa Maria, the Central Pico, Faial, São Jorge, Graciosa and Terceira, and the Western Flores and Corvo. Corvo, with 17 km^2, is the smallest and the northernmost island which lies approximately at the same distance from the Iberian Peninsula and Newfoundland in Canada (Borges and Brown 1999; Borges et al. 2008). São Miguel, with 757 km^2, is the largest island and together with Terceira and Faial are the most densely populated islands (Borges et al. 2008). Surrounded by the Atlantic Ocean, the Azores experiences a marine temperate climate, influenced by the warm Gulf Stream, with high humidity, high precipitation and persistent wind (Connor et al. 2012).

The Azores was officially discovered in the 15th century by the Portuguese navigators. Since then, its flora and fauna have been deeply modified by humans, mainly with the introduction of not only many ornamental plant species (e. g. hydrangea, cryptomeria, plecthanthus, azelea) and crops (e. g. pasture, maize, patato, beta saccharin, vineyand, banana, pineapple (Massot 2015) but also animals (e. g. rabbit, house mouse, pig, dog) and grazing animals such as cattle, sheep, goat (Santos et al. 1995), including the honey bee, which was introduced in the 15th century during the colonization of the archipelago (De La Rúa et al., 2006).

According to the data provided by “Secretaria Regional da Agricultura e Ambiente”, in 2014 and 2015, the years in which the sampling was carried out, São Miguel presented the highest levels of beekeeping activities with the largest number of beekeepers and colonies, followed by Terceira and Pico (see supplementary material Table S3). In 2015, colonies of *A. mellifera* were introduced for the first time on Corvo, thus giving rise to beekeeping on this island.

2.2. Sampling

Between July and August of 2014 and 2015, adult honey bee workers were collected from the inner part of 473 hives distributed across 153 apiaries located in eight islands of the Azores Archipelago (Table 1). Until the year 2015 there was no record on beekeeping activity or the existence of the *A. mellifera* in Corvo Island. Honey bees were shipped to the CIMO laboratory in Bragança and then stored in absolute ethanol at -20ºC until DNA extraction and collection of the forewings.

Table 1. Total number of apiaries and colonies sampled in the Azores Archipelago.

Island	Number of apiaries	Number of colonies
Santa Maria	17	56
São Miguel	33	99
Terceira	26	78
Graciosa	7	21
São Jorge	13	37
Faial	19	60
Pico	24	75
Flores	14	47
Total	153	473

2.3. *MtDNA* intergenic tRNA^{leu}-cox2 region

2.3.1. DNA extraction and statistical analyses

The DNA was extracted from the thorax of a single worker per colony using the Ron's Tissue DNA Mini kit (BIORON®). The intergenic tRNA^{leu}-cox2 region was PCR-amplified with the primer pair E2 (5'-GGCAGAATAAGTGCATTG -3'), located close to the 5' end of tRNA^{leu}, and H2 (5'-CAATATCATTGATGACC-3'), located close to the 5' end of cox2. The PCR reaction was conducted according to the conditions detailed in Garnery et al. (1993), with slight modifications. The reaction was performed in 25 µl total volume containing 2 mM of dNTP, 20 mM $MgCl_2$, 2 pM of each primer (E2 and H2), 5 µl of buffer solution (5X Green GoTaq® Flexi Buffer), 0.75 U µl of GoTaq® Flexi DNA Polymerase (Promega®) and finally, 1 µl of DNA. The PCR temperature profile consisted of an initial denaturing step of 5 minutes at 94ºC followed by 35 cycles of 45 seconds at 92º C; 45 seconds at 48º C, 2 minutes at 62º C, and a final extension of 20 minutes at 65 ºC in a T100™ Thermal Cycler BIO-RAD. The sizes of the PCR products were resolved in a 1% agarose gel using as a molecular marker the BenchTop 100bp DNA Ladder (Promega®). The gel images were acquired with the Molecular Imager® Gel Doc™ XR System BIO-RAD and visualized with the software Image Lab™ 2.0. The PCR products were digested with the *Dra*I restriction endonuclease (Promega ®). The

recognition sequence of *Dra*I is 5´…TTT▼AAA…3´ and in the reverse sequence 3´…AAA▲TTT…5´. The fragment size patterns obtained from digesting the PCR products with the *Dra*I were determined on a 4% wide-range agarose gel (Sigma-Aldrich®) and identified following the complete set of restriction fragment size reported to date and sequence data available in GenBank. The haplotypes were named following the recent revision of the nomenclature system proposed by Chávez-Galarza et al. (2017). Due to resolution limitations of the wide-range agarose gel, it was not possible to distinguish haplotypes whose band fragments show similar patterns (Chávez-Galarza et al. 2017), which is the case of variants within C lineage (Franck et al. 2000b). Therefore, many samples of C-lineage ancestry were PCR-amplified and sent for STAB VIDA (Caparica, Portugal) for Sanger sequencing to determine the haplotypes (e. g. C1 and C2) present in the Azores. In addition to the C-lineage haplotypes, two of the 14 colonies exhibiting a similar *Dra*I band were sequenced to determine whether it was A9 (47,783 830 bp), A30 (47,768 815 bp) or both haplotypes that were present in the Azores. Haplotypes exhibiting novel band patterns were also sequenced for later deposition in GenBank. Sequences were checked and aligned with MEGA 6.06 (Tamura et al. 2013). Genetic diversity parameters were calculated using the software GENEALEX 6.5 (Peakall and Smouse 2012).

2.4. Geometric Morphometrics of the wings

2.4.1. Reference samples

The reference collection comprised forewing samples of the Iberian honey bee, *A. m. iberiensis*, and the two C-lineage subspecies preferred by beekeepers worldwide, *A. m. ligustica* and *A. m. carnica.* The Iberian honey bee a collection of 711 colonies sampled from 237 apiaries distributed across three transects (Chávez et al. 2013): Atlantic transect, Central transect, and Mediterranean transect with sampling details on supplementary material (Figure S1). The C-lineage reference collection consisted in 11 colonies of *A. m. ligustica* and 15 colonies of *A. m. carnica* obtained from the Morphometric Bee Data Bank in Oberursel, Germany (Ruttner's collection, Ruttner 1988). These two C-lineage subspecies were selected because they have seemingly been introduced in other Macaronesian archipelagos, such as in the Canaries (De la Rúa et al. 2006; Muñoz et al. 2013).

2.4.2. Acquisition of forewing data

The right forewing of 5 workers per colony were carefully detached at their base with forceps and mounted in a microscope slide and photographed with a digital

camera attached to a stereomicroscope. Except for the two C-lineage samples, for which the photos had been taken by others, all the forewings were mounted and photographed purposefully for this study. The subsequent analyses were performed using the average of 5 forewings per colony to eliminate intraspecific variation or differences between individuals or measurement errors.

2.4.3 Statistical analyses

The first step of the statistical analysis was to create a TPS file from the images of forewings produced by the software tpsUtil version 1.58 (Rohlf 2004). Subsequently, 19 homologous landmarks were manually plotted at the junction of the veins in the right forewing (Figure 5) by using the software tpsDig version 2.17 (Rohlf 2013). In order to establish a homology among all the specimens, the anatomical landmarks were plotted in the same order in all the wings (Figure 5). Landmarks of all specimens included in this study (Azorean and reference) were plotted by the same person to minimize measurement errors produced by different operators.

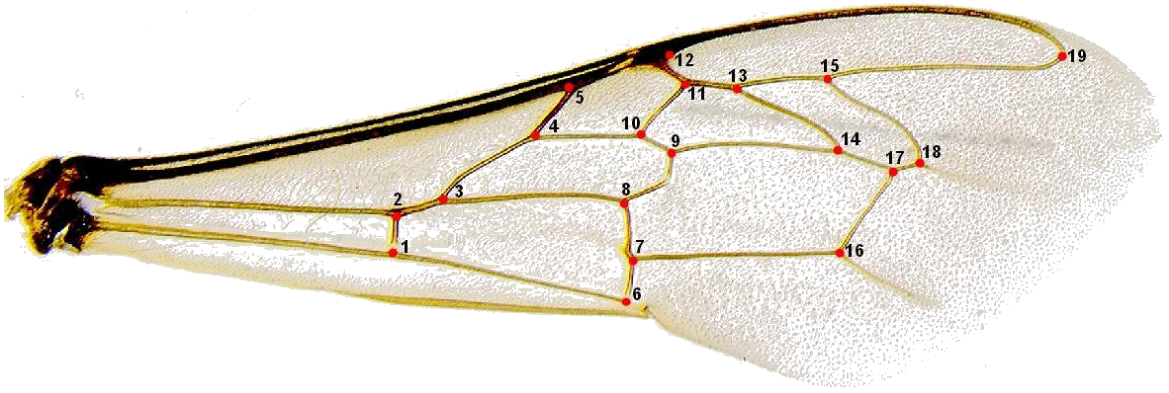

Figure 5. Right forewing of an *Apis mellifera* worker showing the 19 landmarks.

Specimens with missing landmarks were eliminated from the final dataset, resulting in a total of 5925 honey bee forewings, including the reference samples (Table 2).

Table 2. Number of honey bee forewings used through the study.

Samples	Number of forewings
Azores	2287
A. m. iberiensis	3518
A. m. ligustica	55
A. m. carnica	75
Total	5 925

The contribution of each landmark to the observed variation was calculated using the software tpsRelw v1.65 (Rohlf 2005). The software used in the present study for geometric morphometry is freely available at http://life.bio.sunysb.edu/morph/soft-utility.html. All the ensuing analyses were performed using the software MorphoJ v1.06a (Klingerberg 2011). The wings were aligned according to the Procrustes methodology: wing images were scaled, and distances between homologous landmarks were minimized by translations and rotations to achieve an optimal alignment landmark (Bookstein 1991; Zelditch et al. 2004). This procedure eliminated variations caused by size, position and orientation of the specimens (Klingerberg 2011), producing the points of biological shape variation. A global dataset with raw Cartesian coordinates X and Y of each anatomical landmark was generated, which allowed quantifying variation among aligned wings. Shape differences between the Azorean and reference samples resulting from the Procrustes superimposition were investigated through multivariate statistical analyses, which included Principal Component Analysis (PCA), Canonical Variate Analysis (CVA), and Discriminant Function Analysis (DFA), also obtained with MorphoJ v1.06a (Klingerberg 2011). CVA and DFA generate Mahalanobis and Procrustes distances. PCA allow to discriminate variation within a sample and to characterize the main features of shape variation while CVA was used to find the shape features that allow distinguishing among groups. DFA examines the separation between pairs of groups that are known a priori. Given that DFA tends to over-estimate the separation between groups, the reliability of the discrimination can be assessed with the cross-validation test. Corrected classification rates were estimated using the DFA and cross-validation test data. DFA is similar to CVA; however, while DFA is useful for comparisons with specific groups CVA is indicated for general analysis of group structure. The Mahalanobis and Procrustes distances were obtained with the CVA and DFA. While Procrustes distance measures the absolute magnitude of shape variation, Mahalanobis distance provides an indication of how unusual is an individual compared to the others in the sample. The Mahalanobis and Procrustes distances among populations were used in the construction of a neighbor-joining distances dendrogram of morphological proximity implemented in the online software T-REX (Boc et al. 2012).

3. Results

3.1. Mitochondrial DNA

The *Dra*I test detected 13 different haplotypes belonging to three different evolutionary lineages (A, M, C) among the 473 colonies sampled in the Azores (Table 3). These haplotypes represent eight different length patterns resulting from the combination of the different forms of the P element and number of repeats of the Q element. Of the 13 haplotypes, 10 are of African ancestry representing three African sub-lineages (A_I, A_{II}, A_{III}).

Table 3. RFLP length patterns deduced from the *Dra*I test for the 13 haplotypes found in the Azores Archipelago.

Lineage	Sub-lineage	Haplotype	Sequence	RFLP fragment sizes
A	A_I	A1	P_0Q	47/108/483 638
	A_{II}	A10'	P_0QQQ	47/974 1021
		A8	P_0Q	47/591 638
	A_{III}	A14	P_1QQ	47/93/190/484 815
		A14'	P_1QQQ	47/93/191/191/482 1004
		A16'	P_1QQQ	47/93/867 1007
		A20	P_1Q	47/93/483 623
		A30	P_1QQ	47/768 815
		A34	P_1QQ	47/93/224/482 846
		A64'	P_1QQQ	47/93/191/191/145/315 966
M		M7	PQQ	47/95/65/131/65/422 825
		M70	PQQ	145/65/131/65/422 828
C		C	Q	47/41/64/420 572

Geographical distribution and frequencies of the 13 haplotypes detected in the Azores is shown in Figure 6 and Table 4, respectively. Haplotypes belonging to the African lineage were predominant in the Azores (309 in 473, 65.33 %). The majority of these belong to the African sub-lineage A_{III} (295 in 473, 62.37 %). Of the seven haplotypes of sub-lineage A_{III} ancestry (A14, A14', A16', A20, A30, A34, A64'; Table 3 and 4), haplotype A64' was found for the first time. Following the nomenclature recently

revised by Chavez-Galarza et al. (2017), this novel haplotype was named A64' and was added a " ' " symbol due to the presence of three Q elements (Table 3). Moreover, A64' shows the presence of the P_1 element, which is characteristic of the sub-lineage A_{III}. This haplotype was only found in Santa Maria at a low frequency (3 in 56, 0.054; Table 4). Haplotype A14 is the most widespread, being present in all islands, with frequencies ranging from 0.107 (8 in 75) in Pico to 0.987 (77 in 78) in Terceira. Santa Maria presents a haplotype distribution that differs from the remaining islands due to the high frequency of the haplotype A14' (40 in 56, 0.714), which is also present in Faial (6 in 60, 0.100) and with a low frequency in São Miguel (2 in 99, 0.020). Of the 14 samples exhibiting a RFLP pattern ~47/>750 bp, two were sequenced to confirm whether they were A9 (47/783 bp, sub-lineage A_{II}) or A30 (47/768 bp, sub-lineage A_{III}). Sequence data (see supplementary material Figure S2) revealed the presence of the element P1, characteristic of sub-lineage A_{III}, instead of P0, from sub-lineage A_{II}. Therefore, the 14 samples exhibiting that RFLP pattern were assigned to haplotype A30, which was detected only in Santa Maria.

Haplotype A1 was the only representative of sub-lineage A_I and was found in Faial at a very low frequency (2 in 60, 0.033). Sub-lineage A_{II} had two haplotypes: A8 in Santa Maria (3 in 56, 0.054) and A10' in São Miguel (9 in 99, 0.091). Regarding the western European lineage (M), there were only two haplotypes (M7 and M70) among the colonies surveyed. Haplotype M7 was the most frequent and was observed in four colonies of Graciosa (4 in 21, 0.190) and one colony in Terceira (1 in 78, 0.013) whereas M70 was only present in São Miguel in two colonies (2 in 99, 0.020). The eastern European lineage (C) was observed in six of the eight islands surveyed (Figure 6 and 7). Terceira and São Jorge were free of C-lineage haplotypes. High frequencies of C haplotypes were observed in Pico (67 in 75, 0.893) and Faial (45 in 60, 0.750), which contrasted with the low frequencies observed in Flores (4 in 47, 0.085) and Santa Maria (1 in 56, 0.018). Seven private haplotypes from the three African sub-lineages and the M lineage were found in the Azores. São Miguel had the highest number of private haplotypes: two from sub-lineage A_{III} (A34 and A30), one from sub-lineage A_{II} (A10'), and one from lineage M (M70). In Santa Maria, there were two private haplotypes: one from sub-lineage A_{III} (A64') and one from sub-lineage A_{II} (A8). In Faial, there was only one private haplotype from sub-lineage A_I (A8). Except for A30 in São Miguel (0.141), all private haplotypes were detected at low frequencies (< 0.1).

The diversity measures shown in Table 4 indicate that São Miguel carries the greatest genetic diversity (0.786), the highest number of haplotypes (8), and the highest effective number of haplotypes (4). Terceira and São Jorge had the lowest

levels of genetic diversity, with 0.026 and 0.054 respectively, and only with one effective haplotype.

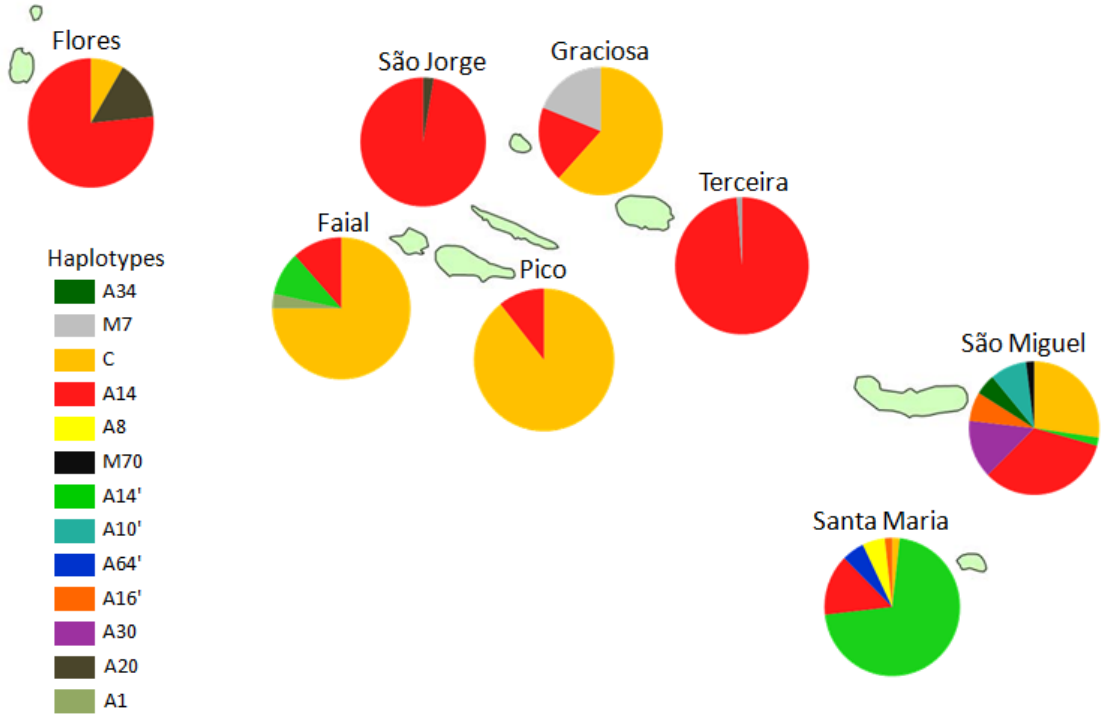

Figure 6. Map with the distribution of the different haplotypes identified with the *Dra*I test in the eight islands surveyed in the Azores Archipelago: Santa Maria (n=56); São Miguel (n=99); Terceira (n=78); Pico (n=67); Faial (n=60); São Jorge (n=37); Flores (n=47); Graciosa (n=21).

Table 4. Distribution of the different haplotypes (number/frequency) and diversity measures for each of the eight islands sampled in the Azores. Haplotypes are grouped by lineage (A, M, C) and African sub-lineage (A_I, A_{II}, A_{III}).

Lineage	Haplotype	Faial	Flores	Graciosa	Pico	São Jorge	Santa Maria	São Miguel	Terceira	Total
A_I	A1	2/0.033								2
	Total									**2**
A_{II}	A10'							9/0.091		9
	A8						3/0.054			3
	Total									**12**
A_{III}	A14	7/0.117	36/0.766	4/0.190	8/0.107	36/0.973	8/0.143	33/0.333	77/0.987	209
	A14'	6/0.100					40/0.714	2/0.020		48
	A16'						1/0.018	7/0.071		8
	A20		7/0.149			1/0.027				8
	A30							14/0.141		14
	A34							5/0.051		5
	A64'						3/0.054			3
	Total									**295**
M	M7			4/0.190					1/0.013	5
	M70							2/0.020		2
	Total									**7**
C	C	45/0.750	4/0.085	13/0.619	67/0.893		1/0.018	27/0.273		**157**
	Total									
	N[a]	60	47	21	75	37	56	99	78	
	Na[b]	4	3	3	2	2	6	8	2	
	Ne[c]	1.703	1.623	2.194	1.235	1.056	1.862	4.502	1.026	
	Priv[d]	1	0	0	0	0	2	4	0	
	uh[e]	0.420	0.392	0.571	0.193	0.054	0.471	0.786	0.026	

[a]Sample size; [b]total number of haplotypes; [c]effective number of haplotypes; [d]number of private haplotypes; [e]unbiased genetic diversity

The representation of the haplotypes by lineage and African sub-lineage shown in Figure 7 highlights the predominance of lineages A and C in the Azores. São Jorge (100%), Terceira (99%), Santa Maria (93%), Flores (91%), and São Miguel (62%) carried mostly A-lineage whereas Pico (89%), Faial (75%), and Graciosa (62%), had large proportions of C-lineage. Lineage M was only present in Graciosa. Regarding representation by sub-lineage, A_{III} was the most frequent across the Azores with São Jorge holding 100% of the colonies. In contrast, sub-lineages A_{I} and A_{II} were observed in low percentages and only in Faial (A_{I}: 3%), Santa Maria (A_{II}: 5%), and São Miguel (A_{II} 9%).

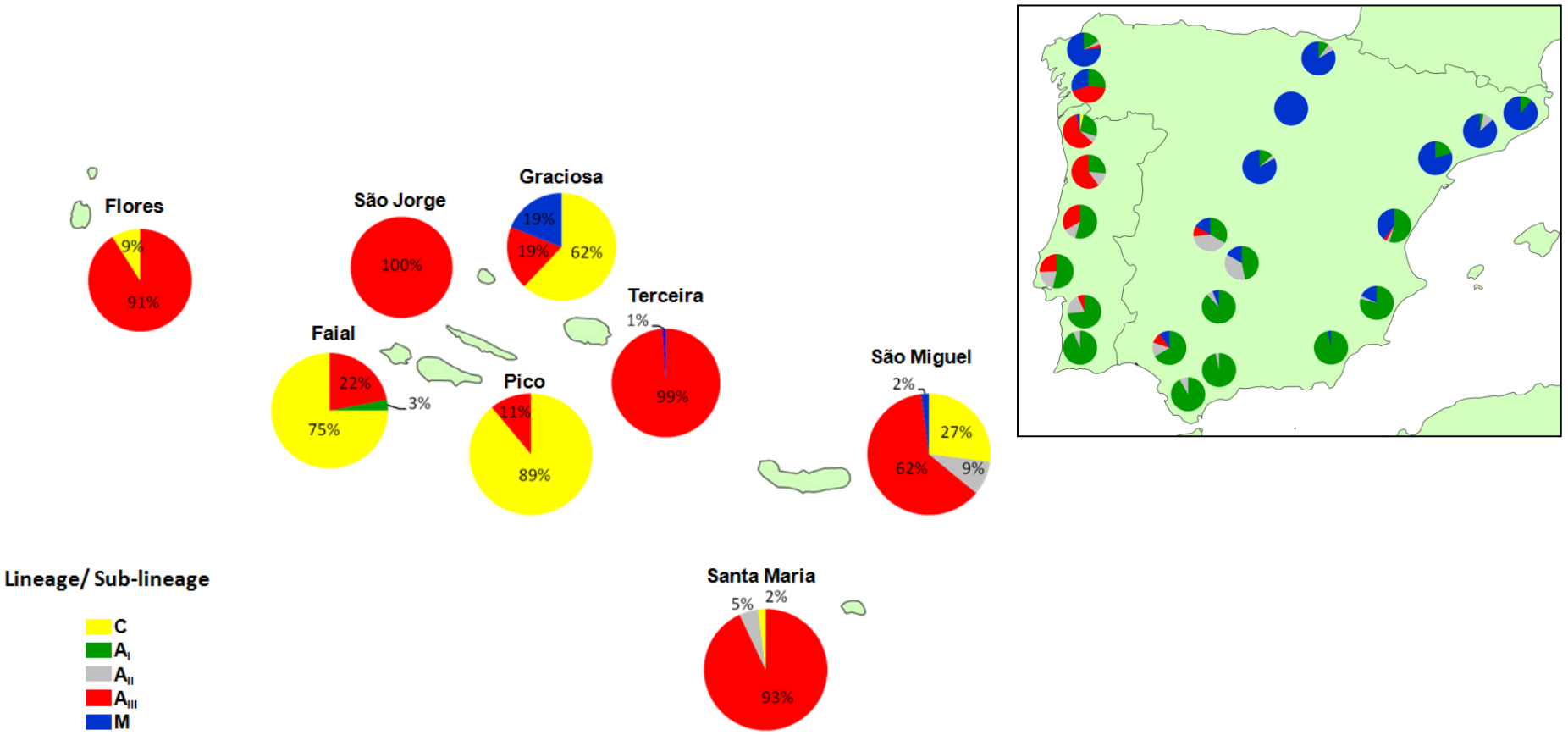

Figure 7. Map w ith the geographical distribution of the different lineages (M, C, A) and African sub-lineages (A_I, A_{II}, A_{III}) observed in the eight islands from the Azores archipelago. An additional map with the different lineages (M, C, A) and African sub-lineages (A_I, A_{II}, A_{III}) obtained for the Iberian Peninsula by Chávez-Galarza et al. (2017) is represented for comparison purposes.

In the 69 individuals of C-lineage ancestry that were sequenced, three different haplotypes were detected: C1, C2j, and C3b (see supplementary materials Figure S2). The three haplotypes differ from each other in one base pair (C1: 572 bp, C2j: 571 bp, C3b: 570 bp). The C1 haplotype was carried by the majority (36) of the sequenced individuals, followed by C2j (29), and C3b (13). Sequence data confirmed that at least three different C-lineage haplotypes are present in Pico and Faial (Table 5).

Table 5. Distribution of C-lineage haplotypes of the 69 sequenced Individuals across the Azores.

Island	Lineage C	Number of sequences	C1	C2j	C3b
Faial	45	15	5	1	9
Flores	4	3	-	3	-
Graciosa	13	10	5	5	-
Pico	67	19	11	4	4
São Miguel	27	21	14	7	-
Santa Maria	1	-	-	-	-
Total	157	68	35	20	13

In this study, the *Dra*I results for 14 colonies which presented the same band pattern raise the doubt if the haplotypes A30 or A9, however sequence data for two of those colonies (see supplementary material Figure S2) reveal the presence of the haplotype A30.

The availability of mtDNA data from previous studies years (De La Rúa et al. 2006, Muñoz et al. 2013) allowed inferring about temporal changes for the island of São Miguel (Figure 8). A decrease of the levels of C-lineage is verified for the sampling carried out in the present study when comparing to previous years. Contrary, a highest percentage of sub-lineage A_{III} is observed for the present study when compared to the other years if sampling.

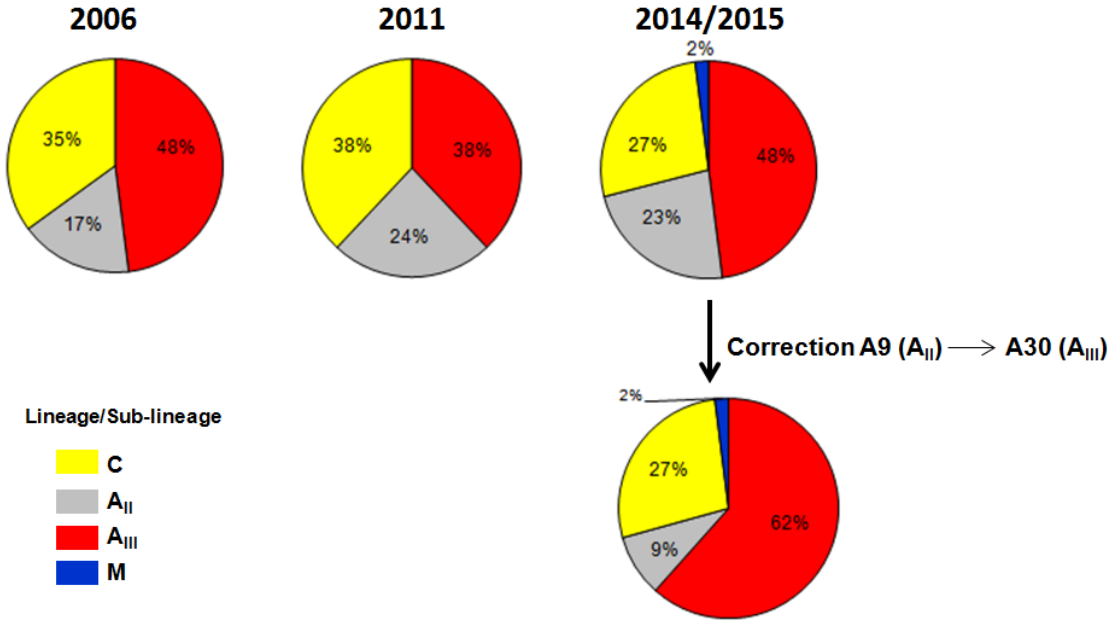

Figure 8. Pie charts showing the present (sampling in 2014/2015) mtDNA survey of honey bees from São Miguel and data from previous studies (sampling in 2006 and 2011) (De La Rúa et al. 2006; Muñoz et al. 2013). The distribution of lineages M and C and African sub-lineages A_I, A_{II}, A_{III} is shown. The pie chart in the line below takes into account the correction of haplotype A9, now A30.

3.2. Wing geometric morphometrics

After plotting the 19 anatomical landmarks in the veins intersection of the wings from the 2287 workers representing the 473 Azorean colonies, a Procrustes superimposition was performed in MorphoJ (Klingerberg 2011). In general, the points showing the Procrustes residuals exhibit a circular distribution around the mean landmark position, as shown in Figure 9 for many landmarks. Contrary to the landmark 8, which shows a tight distribution, landmark 19 exhibited a great dispersion of the individual configurations. The differences observed in the distribution of the individual configurations around the landmarks might be due to random differences between individuals and/or to variation produced by the operator when marking the landmarks. The greater dispersion of points observed for landmark 19 is possibly due to the inherent difficulty of being a semi-landmark. Some individual configurations exhibit an elliptic form around the mean landmark position, as shown by landmarks 3, 4, 10, 9 and, more pronounced for landmarks 14 and 15. This particular variation can be attributed to possible traits under selection (T.Francoy, pers.comm.).

.

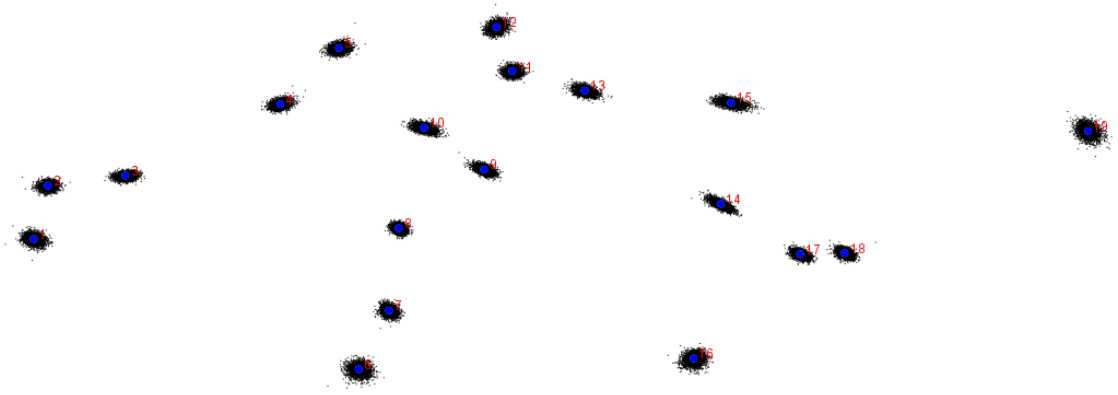

Figure 9. Procrustes residuals of the 19 landmarks in the veins intersections of the honey bee wings from the Azores Archipelago. The blue dots represent the mean landmark position and the small black dots forming a cloud represent the landmark positions for individual configurations in the sample. The number of each landmark is indicated by the red numbers.

The landmarks that contributed the most to the observed variation were landmarks 11 and 17 with 20.07% and 16.50%, respectively (Table 6).

Table 6. Relative contributions of each anatomical landmark for the variation observed in the wings of honey bee from the Azores Archipelago.

Landmark	Variance (%)
1	1.95
2	4.35
3	1.87
4	2.42
5	2.51
6	2.34
7	7.76
8	5.37
9	4.22
10	5.37
11	20.07
12	7.80
13	3.75
14	3.25
15	0.92
16	0.21
17	16.50
18	9.19
19	0.041

After computing the covariance matrix of the Procrustes shape coordinates from the 19 anatomical landmarks, a Principal Component Analysis (PCA) was performed. The PCA allows observing the variation within the different islands and characterizing the principal features of shape variation. The graphs produced by the PCA allow illustrating the shape changes, as is the case of the transformation grids (Figures 10A and 10B) and the wireframe graph (Figures 10C and 10D). For both graph types, it is possible to observe the main landmarks responsible for the variation, although other casual sources of variation can also be observed. The transformation grid shows the shape as a deformation of a rectangular grid using the thin-plate spline. Landmarks exhibiting a greater variation will produce a more accentuated deformation in the grid, as is the case of landmarks 12, 14, 15, and 16 in the PC1 (Figure 10A and 10C). Landmarks 6 and 19 contributed to a more accentuated variation in the PC2 (Figure 10B and 10D).

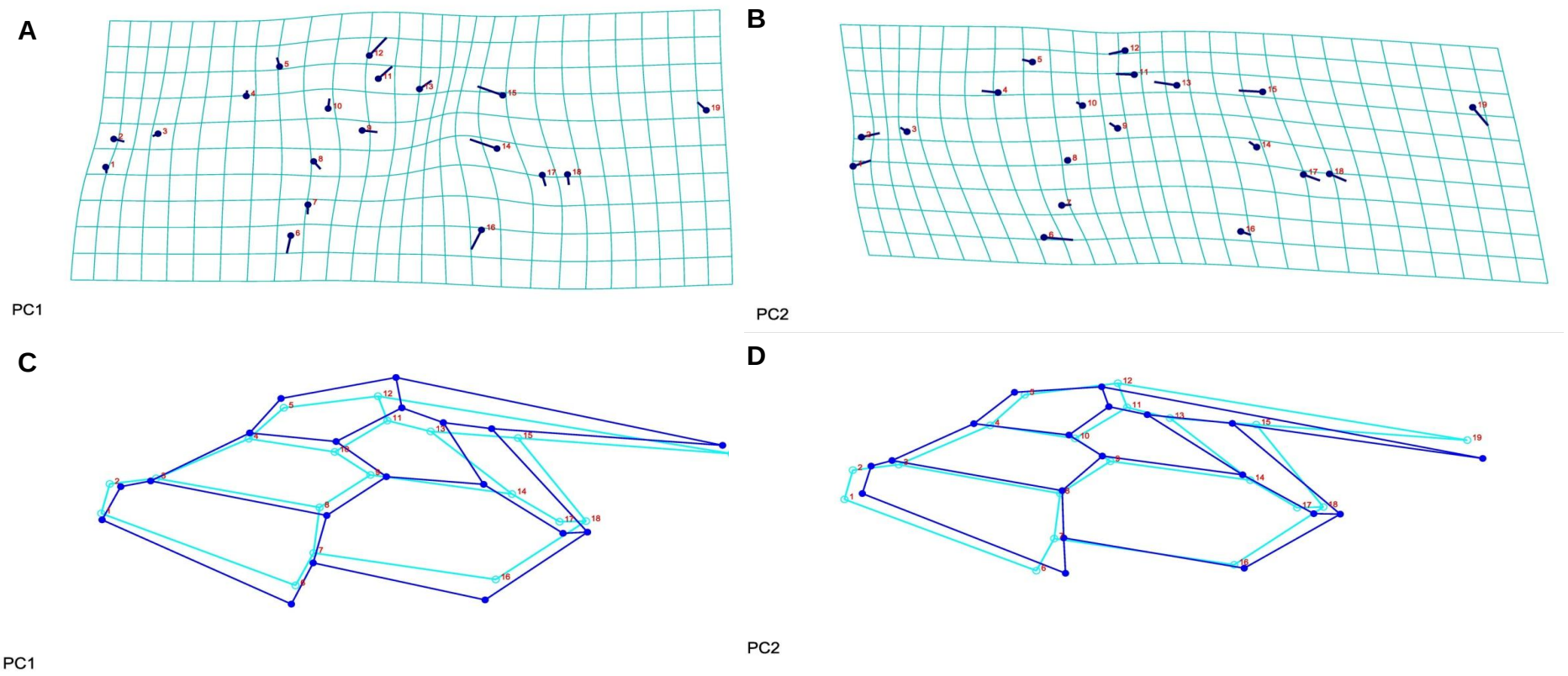

Figure 10. Graphs illustrating the shape variation in the honey bee wings from the Azores Archipelago. Deformation grids of PC1 (A) and PC2 (B) showing the shifts of landmark positions with straight lines with a thin-plate spline. Each line starts with a dot at the location of the landmark in the starting shape (mean shape) and the length and direction indicates the movement of the respective landmark until the target shape. Wireframes of the PCA1 (C) and PC2 (D) showing the 19 landmarks connected with straight lines. The overall mean shape, which is shown with a light blue outline and open dots at the positions of the landmarks, is compared with the target shape represented as a dark blue outline with solid dots. The 19 landmarks are represented in red.

The PCA generated 34 measures of relative warps (k = 2n-4, where “k” refers to the total number of relative deformations and “n” represents the number of anatomical landmarks, which is 19) with each PC showing progressively less variance. The first 14 measures are responsible for 91% of the total variance, although the PC1 (17.0%) and PC2 (14.2%) accounted for more than 31% of the total variance (Figure 11).

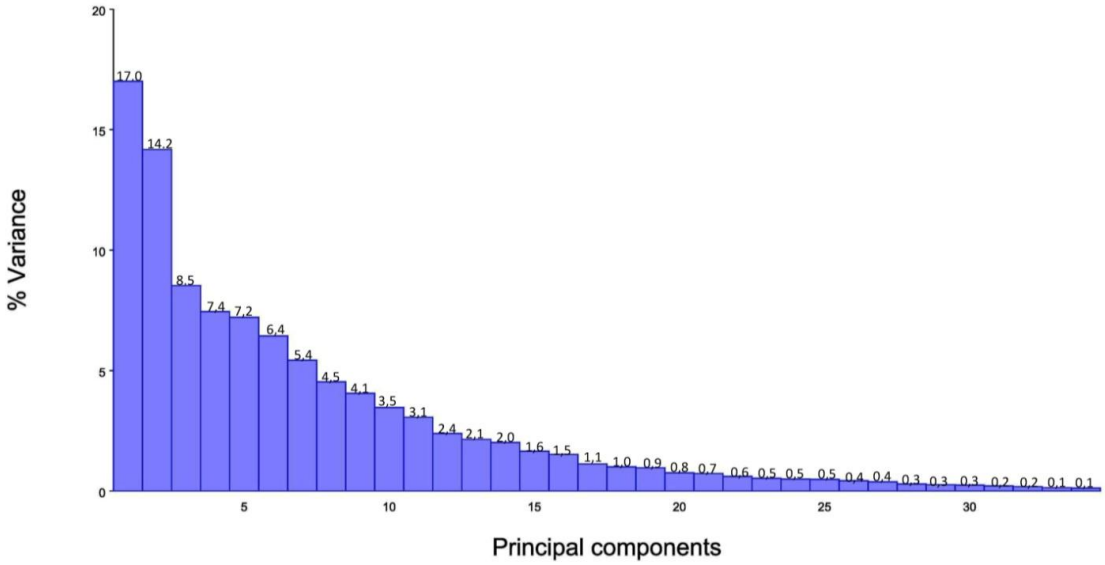

Figure 11. Diagram showing the percentages of total variance for which the 34 PCs account for.

Through the scatter plot obtained using the PC1 and PC2 scores, it is possible to observe differences between populations from each island (Figure 12). Although there is considerable overlap between the specimens of São Jorge, Flores, Pico, and Faial, the specimens of Santa Maria, Graciosa and São Miguel exhibit a different behavior from the remaining islands. The PC1 largely reflects the differences between Santa Maria and Graciosa resulting in two separated clusters. The cluster from São Miguel is situated between that of Santa Maria and Graciosa, showing more similarities with Santa Maria then with Graciosa. The PC2 separated Graciosa from the remaining islands, with the cluster of Graciosa situated in the superior part of the graph.

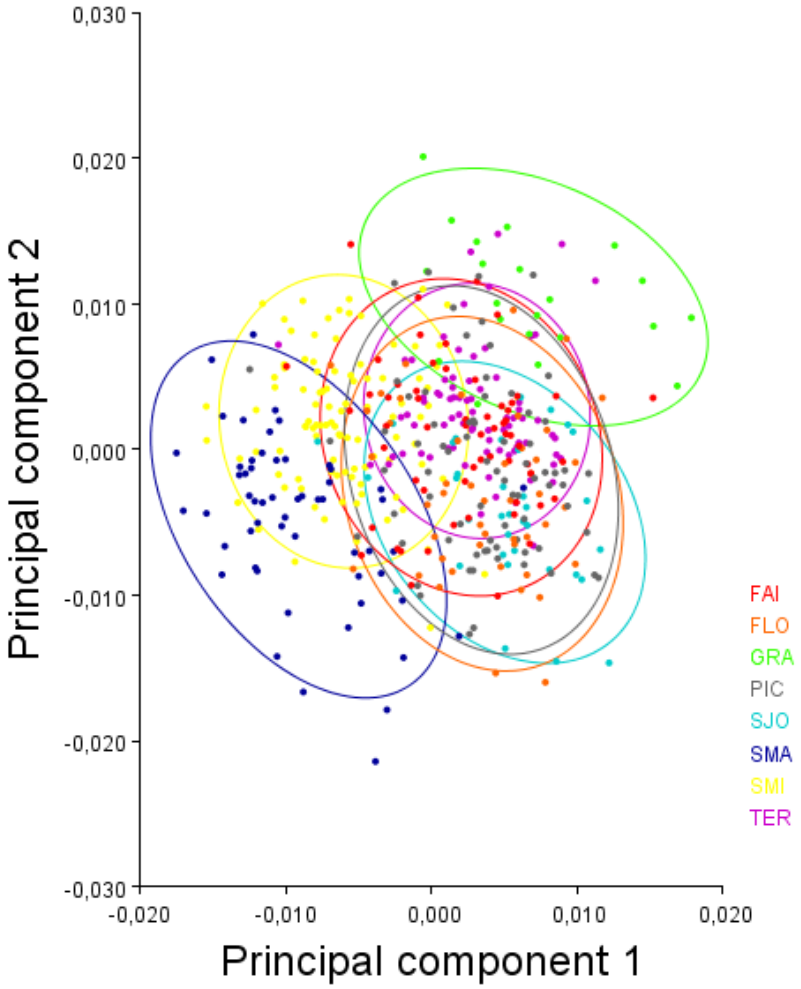

Figure 12. Scatter plot show ing the clusters produced by PCA for the eigh islands of the Azores archipelago analysed in this study. Faial (FAI), Flores (FLO), Graciosa (GRA), Pico (PIC), São Jorge (SJO); Santa Maria (SMA), São Miguel (SMI), and Terceira (TER). Each circle/elipse represents an island. The PC1 is represented in the X axis and PC2 in the Y axis.

Contrary to the PCA, the Canonical Variate Analyses (CVA) is optimized to find differences among the islands. CVA finds the shape features that best distinguish among the different populations. The patterns obtained with the deformation grid (Figure 13A) and the wireframe graph (Figure 13B) show that landmark 14 contributes with the greatest variation, although landmarks 9, 12, 16, 17, and 18 are also important.

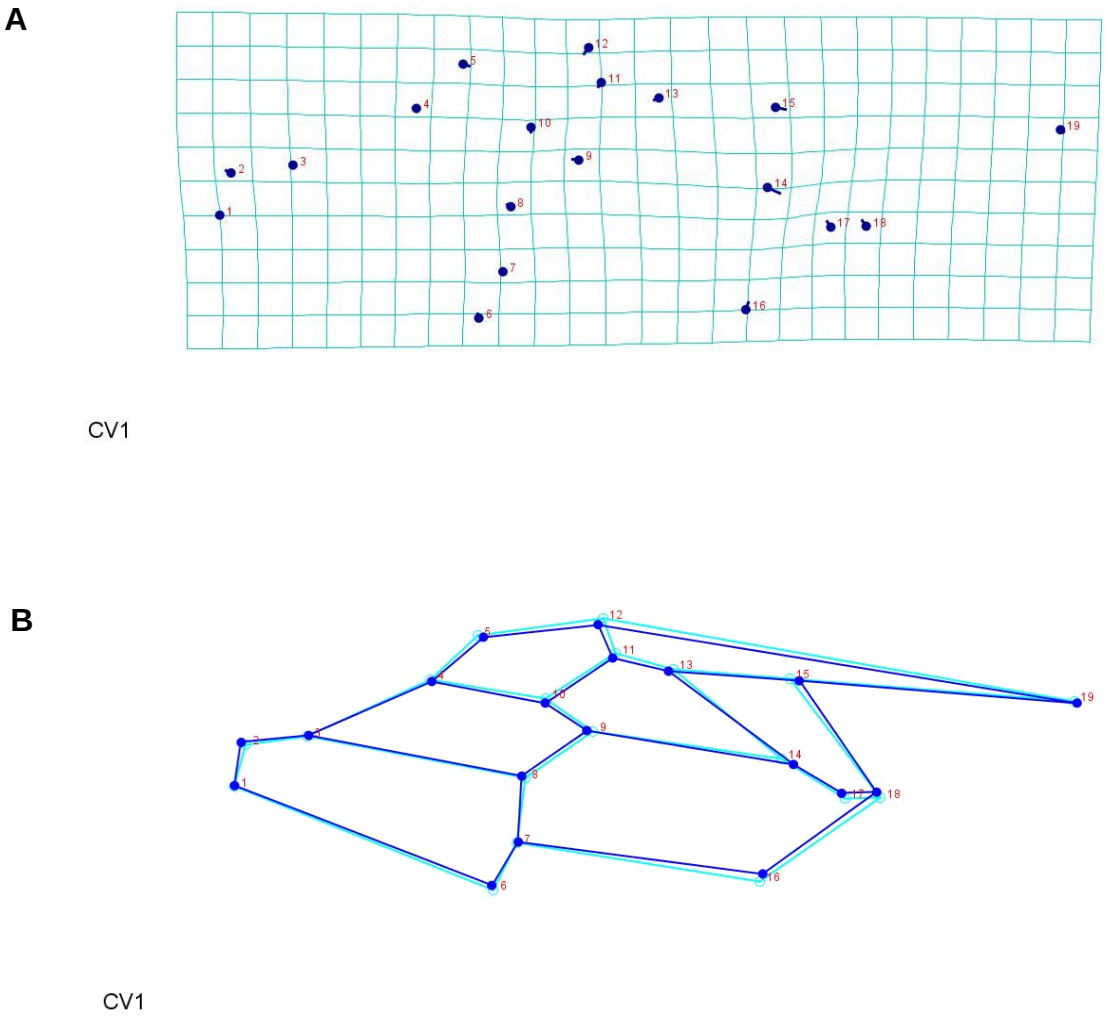

Figure 13. Graphs illustrating the shape variation in the honey bee wings from the Azores Archipelago. (A) Deformation grids of CV1 showing the shifts of landmark positions with straight lines with a thin-plate spline. Each line starts with a dot at the location of the landmark in the starting shape (mean shape) and the length and direction indicate the movement of the respective landmark until the target shape. (B) Wireframes of the CV1 showing the 19 landmarks connected with straight lines. The overall mean shape is shown with a light blue outline and open dots and the positions of the landmarks are compared with the target shape represented as a dark blue outline with solid dots. The 19 landmarks are represented in red.

The shape variation associated with CV1 explained the most discrimination between the populations of the different islands. The first two canonical variates, CV1(43.5%) and CV2 (18.9%), explain the greatest variation (Table 7).

Table 7. Canonical variates generated in the CVA and respective variance.

Canonical Variates	Variance (%)
1	43.466
2	18.926
3	15.379
4	9.759
5	5.823
6	3.852
7	2.696

The first two canonical variates are represented in the scatter plot in Figure 14. Once more, CV1 in the horizontal axis clearly separates the clusters corresponding to the populations from Graciosa and Santa Maria. Moreover, Santa Maria, apart from showing an evident similarity with the cluster of São Miguel, is separated from the remaining islands. In the CV2, the clusters formed by Graciosa and Terceira are situated in the superior part of the graphic exhibiting a close proximity. The clusters formed by Faial, Flores, Pico and São Jorge only exhibit slight differences between them.

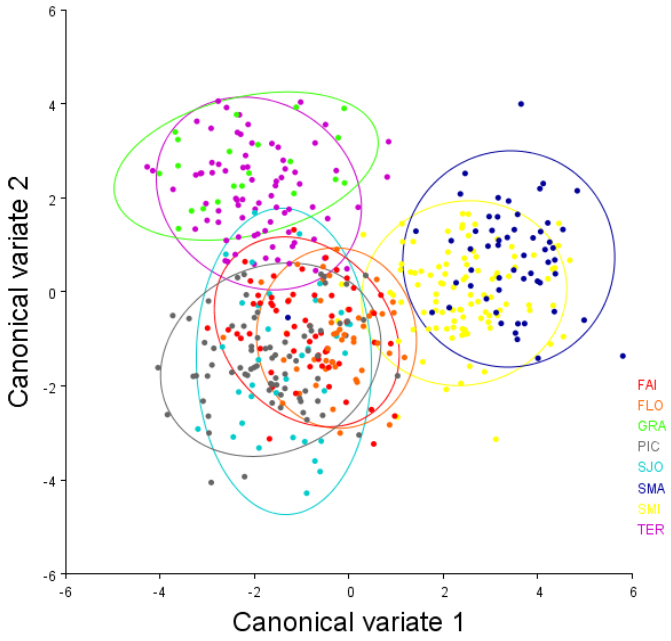

Figure 14. Scatter plot showing the clusters produced by CVA for the eigh islands of the Azores archipelago analysed in this study. Faial (FAI), Flores (FLO), Graciosa (GRA), Pico (PIC), São Jorge (SJO); Santa Maria (SMA), São Miguel (SMI), and Terceira (TER). Each circle/elipse represents an island. The CV1 is represented in the X axis and CV2 in the Y axis.

Mahalanobis and Procrustes distances were calculated using the CVA data and are shown in Table 8. The larger the Mahalanobis distance the greater the differences between islands. The greatest Mahalanobis distances were obtained for comparisons between Graciosa and Santa Maria (7.700), São Jorge (6.6261), São Miguel (6.3421) and Flores (6.1225). In contrast, the lowest values were obtained between Faial and Pico (2.2264). The Procrustes distance, which measures the absolute magnitude of shape variation, shows a similar trend, although at a lower magnitude (Table 8).

Table 8. Procrustes (below the diagonal) and Mahalanobis (above the diagonal) distances between populations, obtained with the CVA data for the wings of honey bee from the eight islands of the Azores Archipelago. All distances were highly significant (*P-value <0.0001)*.

Mahalanobis / Procrustes	Faial	Flores	Graciosa	Pico	São Jorge	Santa Maria	São Miguel	Terceira
Faial	-	3.3197	5.2926	2.2264	3.2558	5.1994	4.0439	3.7013
Flores	0.0082	-	6.1225	3.9902	3.9921	5.1395	4.6538	4.6211
Graciosa	0.0153	0.0164	-	5.6083	6.6261	7.7007	6.3421	4.8892
Pico	0.0064	0.0090	0.0158	-	3.0972	6.0083	4.6965	3.9943
São Jorge	0.0084	0.0082	0.0183	0.0066	-	5.8335	5.1860	4.2095
Santa Maria	0.0146	0.0142	0.0234	0.0149	0.0152	-	3.8311	5.8474
São Miguel	0.0104	0.0126	0.0173	0.0117	0.0140	0.0092	-	5.3221
Terceira	0.0091	0.0098	0.0126	0.0087	0.0103	0.0155	0.0121	-

Whereas CVA produces a comparison between the islands, the DFA (Discriminant Function Analysis) provides pairwise comparisons between specific islands and the reliability of the discrimination, which is assessed by the leave-one-out cross-validation procedure. The DFA correctly identified 98.96% of the colonies, whereas the cross-validation test exhibited 94.39% of accuracy placing the colonies in the respective island. Once more, the differences between Santa Maria and Graciosa are evident in the wireframe graph (Figure 15A) and both DFA histograms (Figure 15B and 15C) clearly show a gap between these two islands.

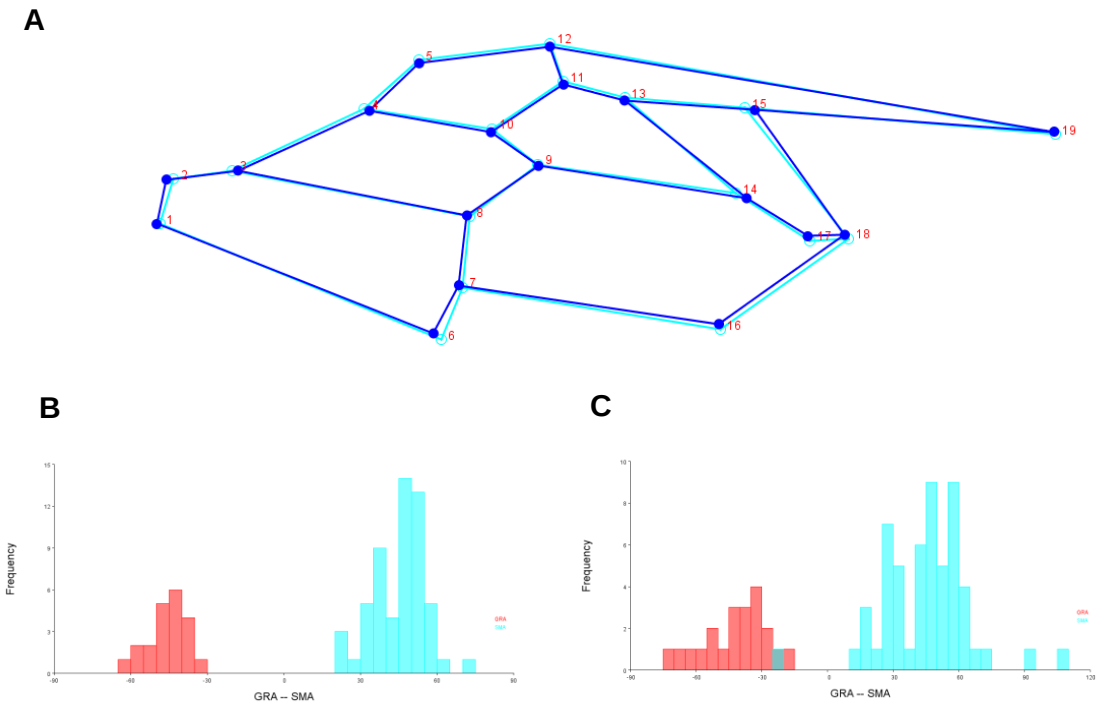

Figure 15. Differences of populations from Graciosa (GRA) and Santa Maria (SMA). (A) The wireframe graph exhibits shape differences, with the light blue outline representing the mean shape and the target shape represented as a dark blue line. (B) Results of the linear discriminant analysis of the differences in wing shape between. (C) Results of the cross-validation test for the discriminant function in wing shape. In the Y axes are represented the frequencies and in the X axes the discriminant function scores.

In contrast, the similarities shared between Faial and Pico are clear according to the wireframe graph (Figure 16A), which only shows differences in a few landmarks. Also, in the histogram graph of the DFA (Figure 16B) and the cross-validation test (Figure 16C) the similarity between the two islands is revealed by the overlapping bars.

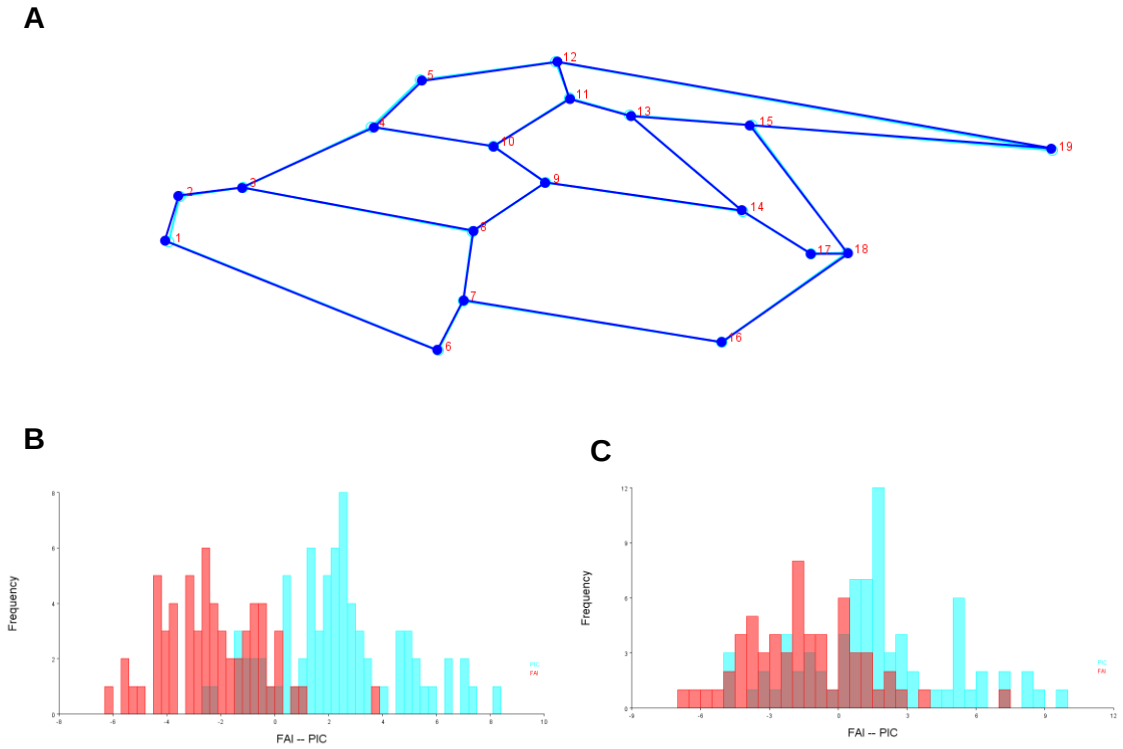

Figure 16. Differences of populations from Faial (FAI) and Pico (PIC). (A) The wireframe graph exhibiting slight shape differences, with the light blue outline representing the mean shape and the target shape represented as a dark blue line. (B) Results of the linear discriminant analysis of the differences in wing shape. (C) Results of the cross-validation test for the discriminant function in the wing shape. In the Y axes are represented the frequencies and in the X axes the discriminant function scores.

The Mahalanobis and Procrustes distances obtained with DFA highlighted once more for the great differences between populations from Graciosa and Santa Maria and the strong affinity of populations from Pico and Faial (Figures 15 and 16) (see supplementary material Table S4). The results obtained for the morphology are not coherent when compared with the geographical distances between islands.

Wing morphometric data of the Azorean samples were then compared with the reference samples from the Iberian Peninsula (*A. m. iberiensis*), the putative origin of the Azorean honey bees, and from Eastern Europe (*A. m. ligustica* and *A. m. carnica*), the origin of putative recent imports. The first PCA explained 19.36% of the variance and reveals a closer relationship of the samples from Graciosa with the C-lineage reference samples of *A. m. carnica* and *A. m. ligustica*, rather than those of *A. m. iberiensis* (Figure 17). The clusters formed by the samples of Flores, Faial, Pico, and Terceira exhibited a subtle overlap with that of *A. m. carnica.* Clusters formed by the samples of Graciosa, *A. m. ligustica* and *A. m. carnica* tend to be located at the top of the scatter plot, contrary to the remaining clusters, as shown by the PC2 which explains 11.95% of total the variance. The cluster of Graciosa was the only one that

overlapped with those of *A. m. ligustica* and *A. m. carnica*. On the other hand, the clusters of São Jorge, Santa Maria and São Miguel were de only populations in the island that did not overlap with *A. m. ligustica* and *A. m. carnica* (Figure 17).

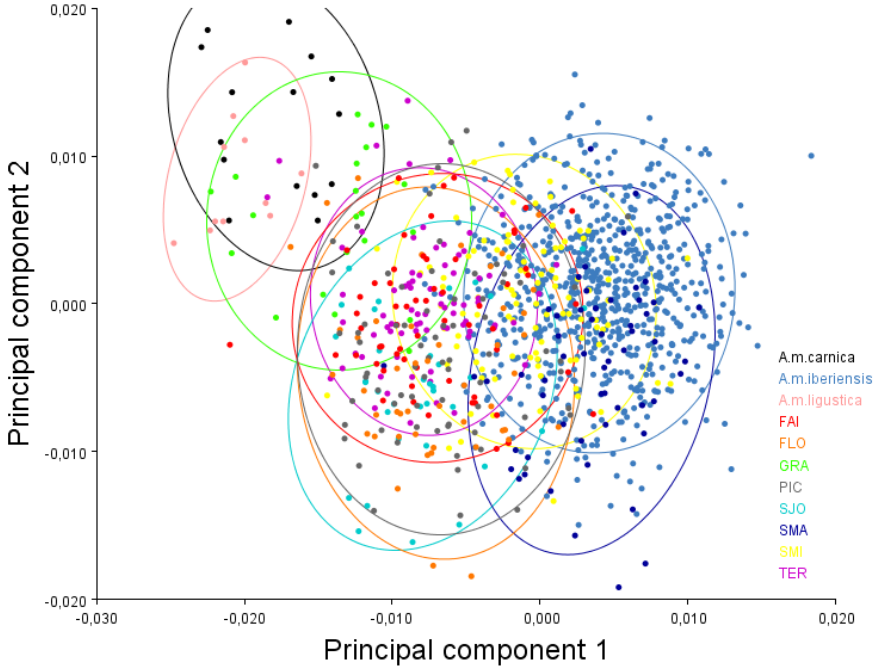

Figure 17. Scatter plot showing the clusters generated by PCA for the eight islands of the Azores archipelago analysed in this study: Faial (FAI), Flores (FLO), Graciosa (GRA), Pico (PIC), São Jorge (SJO), Santa Maria (SMA), São Miguel (SMI), and Terceira (TER). The reference samples of *A. m. ligustica*, *A. m. carnica* and *A. m. iberiensis* were included in the analyses. Each circle/ellipse represents an island or subspecies. The PC1 is represented in the X axis and PC2 in the Y axis.

The results of CVA are shown in Figure 18. Shape changes associated with CV1 (51.27%) explained the most discrimination between samples, along with CV2 (21.42%). Overall, populations from the different islands of Azores were clustered together and overlapped with each other. The only exception was observed for the populations of Graciosa and Santa Maria, which showed again a clear separation (Figure 18). The cluster of Graciosa exhibited a close proximity with those of *A. m. ligustica* and *A. m. carnica,* which were clustered together and clearly separated from the remaining populations. The clusters formed by the samples of São Miguel and Santa Maria exhibited a high degree of overlapping, being both slightly separated from the other islands (Figure 18). Moreover, the cluster of São Miguel was the only one that overlapped with the reference samples from the Iberian Peninsula (Figure 18).

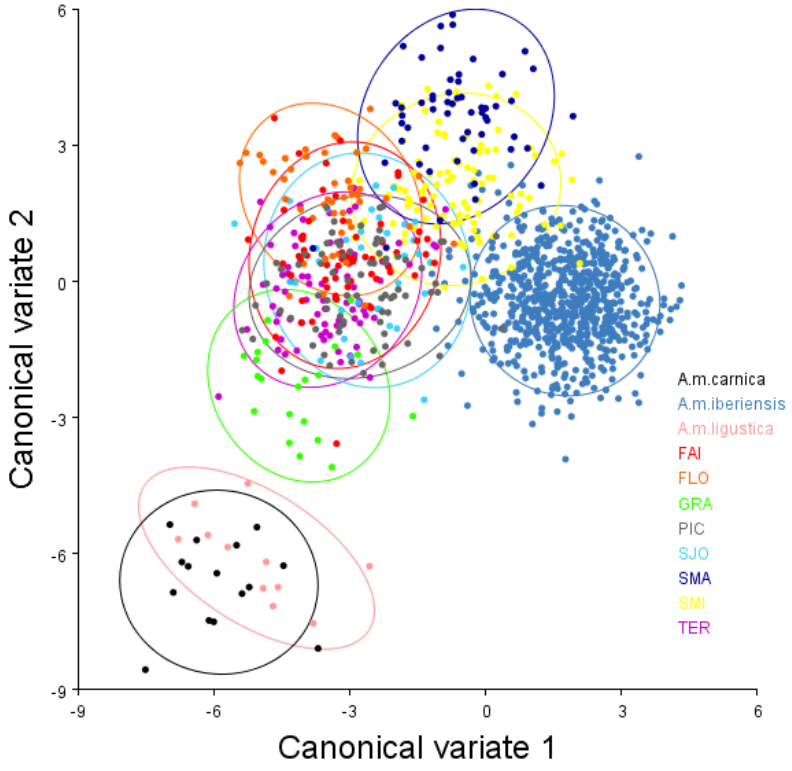

Figure 18. Scatter plot showing the clusters generated by CVA for the eight islands of the Azores archipelago analysed in this study: Faial (FAI), Flores (FLO), Graciosa (GRA), Pico (PIC), São Jorge (SJO), SMA (SMA), São Miguel (SMI), and Terceira (TER). The reference samples of *A. m. ligustica*, *A. m. carnica* and *A. m. iberiensis* were included in the analyses. Each circle/ellipse represents an island. The CV1 is represented in the X axis and CV2 in the Y axis.

Mahalanobis and Procrustes distances obtained with the CVA (Table 9) produced the highest values for Santa Maria when compared with *A. m. carnica* (MD=12.0807, PD=0.0323; *p-value <0.0001*) and *A. m. ligustica* (MD=11.1991, PD=0.0310, *p-value <0.0001*). Faial and Pico were the closest islands (MD=2.1955, PD=0.0064; *p-value <0.0001*) (Table 9). Although the Mahalanobis and Procrustes distances exhibited overall a similar pattern, there were some discrepancies between both measures. While the lowest values of the Mahalanobis distance for the comparisons between *A. m. iberiensis* and the Azorean samples were obtained for São Miguel (MD=4.2072, *p-value <0.0001*) and Pico (MD=4.9759, *p-value <0.0001*), those of the Procrustes distance were obtained for São Miguel (PD=0.0101, *p-value <0.0001*) and Santa Maria (PD=0.0107, *p-value <0.0001*) (Table 9). However, the differences between both distances obtained with the CVA data were subtle.

Table 9. Procrustes (below the diagonal) and Mahalanobis (above the diagonal) distances, obtained with the CVA data for the wings of honey bee from the eight islands of the Azores Archipelago and the reference subspecies *A. m. iberiensis*, *A. m. ligustica,* and *A. m. carnica.* All distances were highly significant (*P-value <0.0001).*

Procrustes \ Mahalanobis	Faial	Flores	Graciosa	Pico	S. Jorge	S. Maria	S. Miguel	Terceira	iberiensis	ligustica	carnica
Faial	-	3.2622	5.4181	2.1955	3.2962	5.2214	4.2011	3.7076	5.2111	8.3124	9.1927
Flores	0.0082	-	6.3198	3.8015	3.8749	5.0265	4.7063	4.6234	6.0359	9.3974	10.1171
Graciosa	0.0153	0.0164	-	5.6919	6.6332	7.9410	6.6823	4.7613	7.1868	6.1817	7.6876
Pico	0.0064	0.0090	0.0158	-	3.0539	5.8509	4.7520	3.9931	4.9759	8.2458	9.0919
S. Jorge	0.0084	0.0082	0.0183	0.0066	-	5.8164	5.2369	4.4417	5.2406	9.8754	9.4621
S. Maria	0.0146	0.0143	0.0234	0.0149	0.0152	-	3.8017	5.8416	5.3246	11.1991	12.0807
S. Miguel	0.0104	0.0127	0.0173	0.0117	0.0140	0.0092	-	5.5255	4.2072	9.7374	10.9452
Terceira	0.0091	0.0098	0.0126	0.0087	0.0103	0.0155	0.0121	-	5.8033	8.2581	9.2194
iberiensis	0.0133	0.0142	0.0203	0.0126	0.0147	0.0107	0.0101	0.0136	-	9.7699	10.7774
ligustica	0.0217	0.0228	0.0134	0.0218	0.0231	0.0310	0.0256	0.0192	0.0272	-	4.0286
carnica	0.0244	0.0261	0.0182	0.0240	0.0260	0.0323	0.0281	0.0219	0.0281	0.0110	-

DFA correctly classified 99.37% of the colonies in the respective island and groups of subspecies. The cross-validation test, with a high percentage of correct classification (94.58%), supports a strong reliability of the clustering. Regarding the DFA, Mahalanobis distances between groups are slightly different from those estimated by CVA, mostly exhibiting higher values (see supplementary material Table S5). However, the results obtained for Procrustes distances in the DFA were virtually the same as in the CVA.

Figure 19 shows the dendrogram of morphological proximity constructed with the Mahalanobis square distances. Three major clusters, which seem to be closely related with the geographical location of the islands, are observed. Samples from São Miguel and Santa Maria were clustered with *A. m. iberiensis*, suggesting a closer proximity between populations of these easternmost islands and those of the Iberian Peninsula. On the other hand, Graciosa was grouped with the C-lineage subspecies *A. m. ligustica* and *A. m. carnica*. The similarity between Pico, Faial, São Jorge, and Flores is clearly evidenced since these islands were grouped together in the same cluster. It was also built a dendrogram with the Procrustes distances and the results largely corroborated with the Mahalanobis distances dendrogram presented in Figure 19.

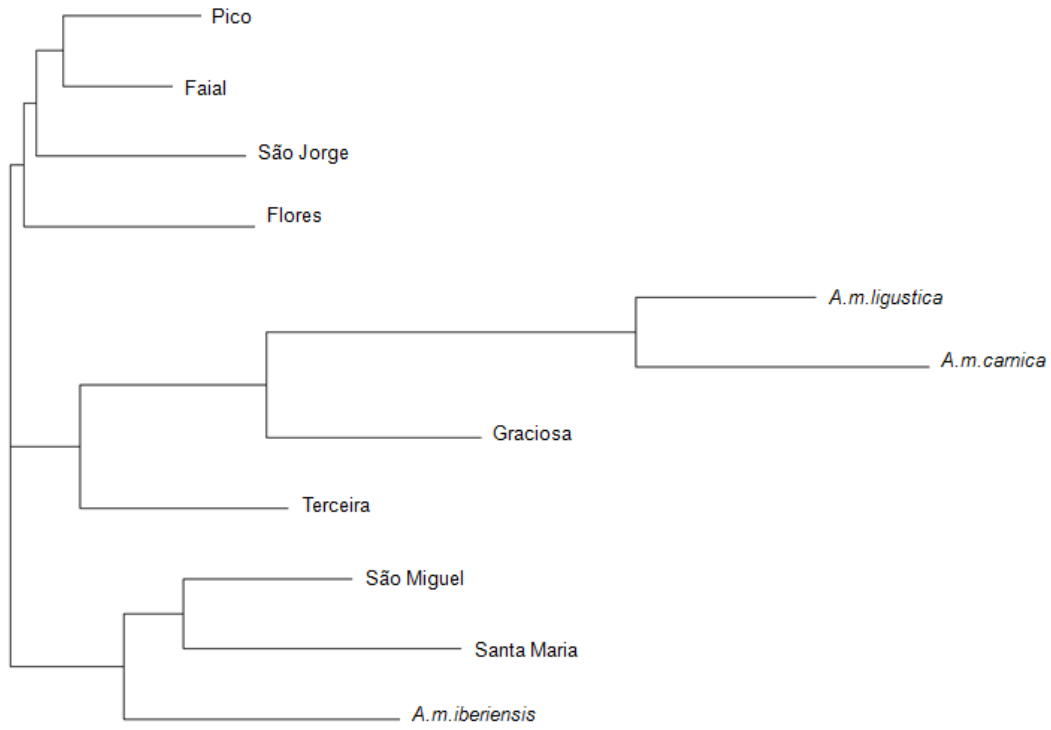

Figure 19. Neighbor-joining dendrogram of morphological proximity between the eight islands of Azores archipelago and the reference samples *A. m. iberiensis*, *A. m. ligustica* and A. *m. carnica* constructed based on the Mahalanobis square distances.

Given that the mtDNA data suggests a close proximity between populations from the Azores and those of continental Portugal, populations from *A. m. iberiensis* previously used, were separated in two groups: one formed by the samples from Portugal and the other from Spain. For the PCA, the two first principal components of the PCA explained 18.44% (PC1) and 11.93% (PC2) of the variation. The cluster containing populations from the Azores exhibited some degree of overlap with the individuals of Portugal and Spain, with a tendency to separate them (Figure 20).

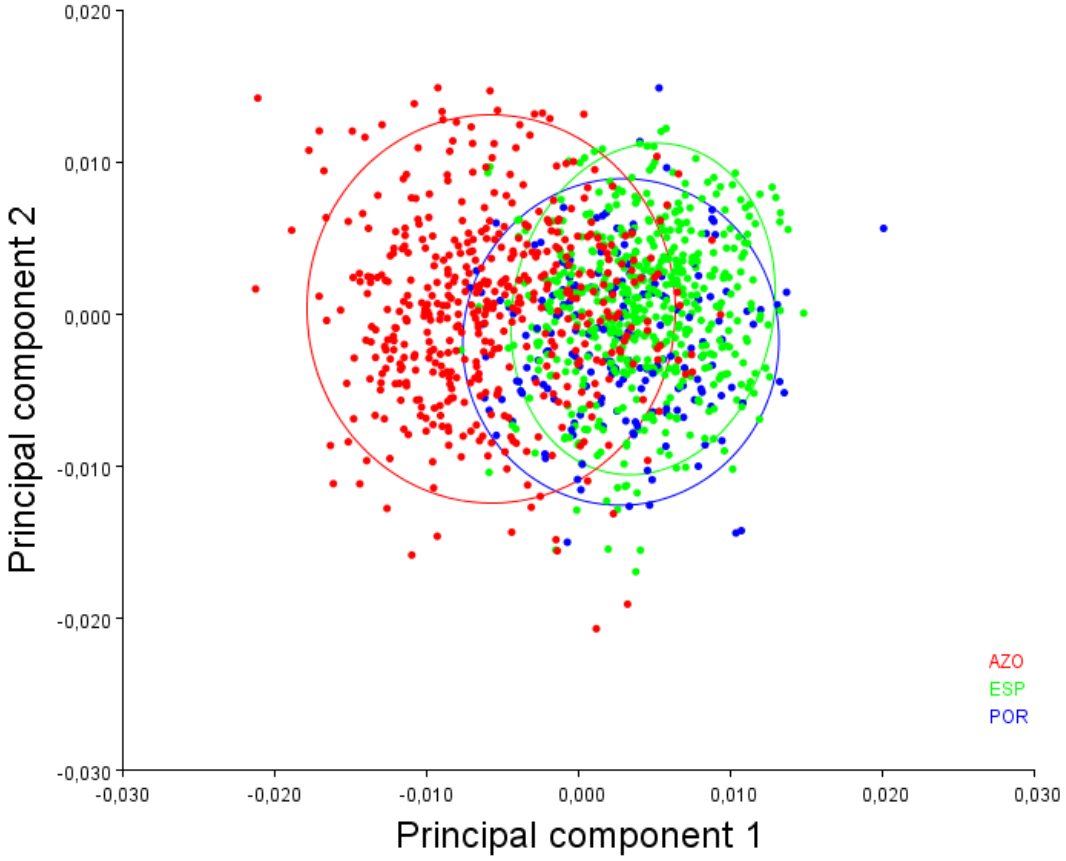

Figure 20. Scatter plot showing the clusters produced by PCA for three populations; Azores (AZO); Spain (ESP), and Portugal (POR). Each circle/ellipse represents a population. PC1 is represented in the X axis and PC2 in the Y axis.

For the CVA, two CV were generated with CV1 accounting for 93.95% of the total variance and CV2 explaining only 6.052% (Figure 21). In fact, in the CV1 axis it is clear the formation of two separated clusters, one comprising the Azorean samples and other those from Spain. Additionally, populations from the Azores exhibited a closer relationship with those from Portugal showing partial overlap.

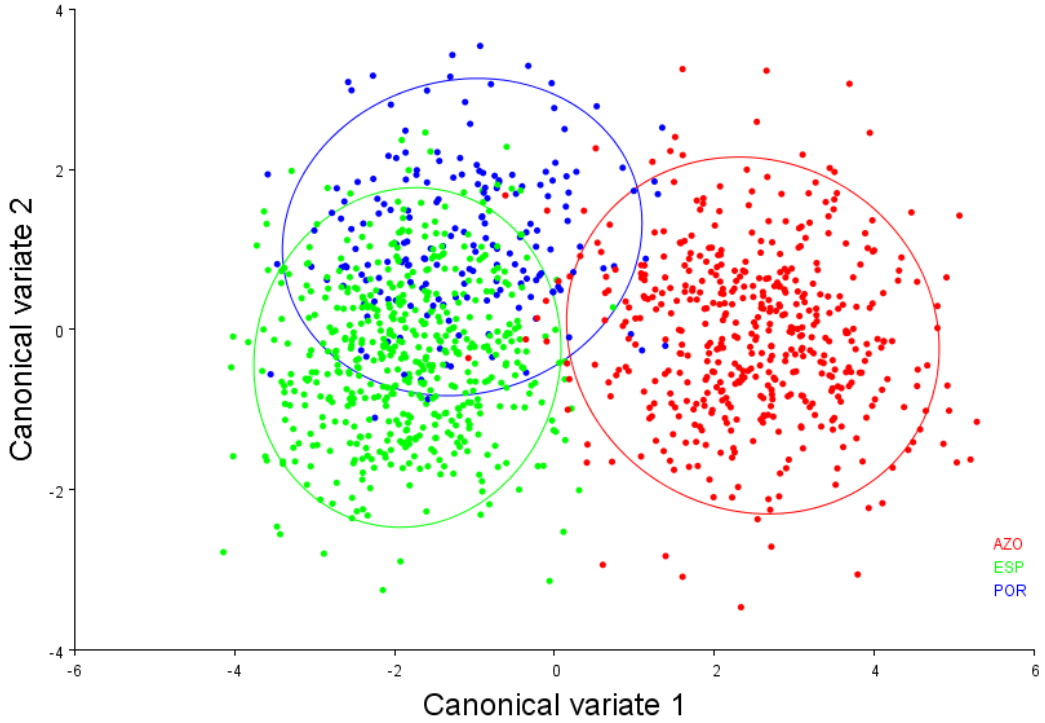

Figure 21. Scatter plot showing the clusters generated by CVA for three populations: Azores (AZO), Spain (ESP), and Portugal (Por). Each circle/ellipse represents one such population. The CV1 is represented in the X axis and CV2 in the Y axis.

Values obtained from the CVA for the Mahalanobis distance confirmed that there is a greater similarity of the populations from the Azores with those of continental Portugal (3.8470, *p-value* <0.0001) than those of Spain (4.3322, *p-value* <0.0001) (Table 10). Although at a smaller amplitude, results obtained for the Procrustes distance also demonstrate a great similarity of populations from the Azores with populations of *A. m. iberiensis* from Portugal (Table 10).

Table 10. Procrustes (below the diagonal) and Mahalanobis (above the diagonal) distances, obtained with the CVA data for the wings of honey bee from the Azores Archipelago and populations from A. m. iberiensis from Portugal and Spain. All distances were highly significant (*P-value <0.0001).*

Mahalanobis / Procrustes	Spain	Portugal	Azores
Spain	-	1.6564	4.3322
Portugal	0.0054	-	3.8470
Azores	0.0110	0.0096	-

The DFA correctly classified the majority of the colonies in the respective Island and country (92.36%) whereas the cross-validation test correctly reassigned 90.43% of the colonies to the respective country and island. Mahalanobis and Procrustes distances obtained with DFA data were the same as those obtained with the CVA (see supplementary material Table S6).

Since all mtDNA lineages and sub-lineages were known for the 473 samples analysed for the Azores and for all the reference samples *A. m. iberiensis*, *A. m. ligustica* and *A. m. carnica*, a CVA was performed using the lineages/sub-lineages to found differences between colonies.

Results obtained for CVA overlapped populations from Graciosa, Faial and Pico with populations from *A. m. ligustica* and *A. m. carnica*, which belong to the evolutionary C-lineage. The remaining populations from the islands exhibited proximity with populations from *A. m. iberiensis* (Figure 22).

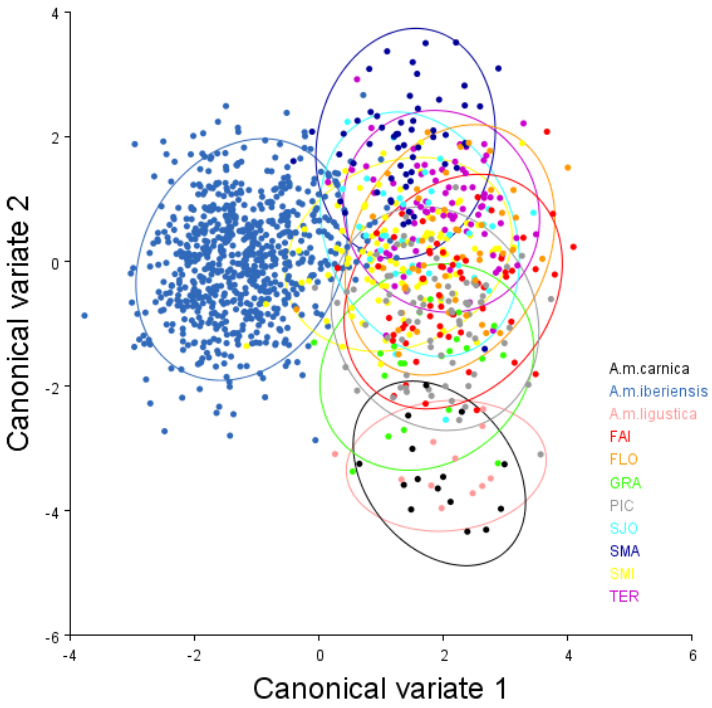

Figure 22. Scatter plot showing the clusters generated by CVA when the variable lineage/sub-lineage were added, for the eight islands of the Azores archipelago analysed in this study: Faial (FAI), Flores (FLO), Graciosa (GRA), Pico (PIC), São Jorge (SJO), SMA (SMA), São Miguel (SMI), and Terceira (TER). The reference samples of *A. m. ligustica*, *A. m. carnica* and *A. m. iberiensis* were included in the analyses. Each circle/ellipse represents an island. The CV1 is represented in the X axis and CV2 in the Y axis.

4. Discussion

4.1. Mitochondrial DNA

This study represents the most comprehensive genetic survey of the honey bee populations from the Azores, with a large number of colonies sampled across the eight islands with beekeeping activity. A total of 473 colonies were analyzed with the *Dra*I test revealing 13 different haplotypes of the mtDNA, including a novel haplotype private to Santa Maria. Moreover, the *Dra*I test also allowed detecting the presence of three divergent evolutionary lineages A, M and C, including the three African sub-lineages A_I, A_{II}, A_{III}.

The African sub-lineage A_{III} was the most distribute lineage and sub-lineage in the Azores. The element P_1 characteristic of the African sub-lineage A_{III} was observed in seven different haplotypes: A14, A14', A16', A20, A30, A34, and A64'. Haplotype A64', private to Santa Maria, was reported here for the first time. This novel haplotype adds to the remarkably high levels of genetic diversity of sub-lineage A_{III}, mainly found in the Atlantic side of the Iberian Peninsula (Pinto et al. 2012, 2013; Chávez-Galarza et al. 2017), but also in the Canaries and Madeira (Muñoz et al. 2013). A14 is the only haplotype that occurs in every Azorean island and it is one of the most widely distributed across Macaronesia, although its frequency is higher in the Azores than in the Canaries (De la Rúa et al. 1998, 2001; Muñoz et al. 2013; Miguel et al. 2016). Along with A14, haplotype A14', with a high frequency in Santa Maria and also present in Faial and São Miguel, was the most common haplotype in Madeira and across the Canary Islands, (De La Rúa et al. 1998, 2001, 2002; Muñoz et al. 2013; Miguel et al. 2016).These two haplotypes were described for the first time by Franck et al. (2001), who also detected A14 in *A. m. adansonii* from Namibia, in addition to the Canaries (Franck et al. 2001). They were found in Portugal and Southern Spain, although at low frequencies (De La Rúa et al. 2002, 2004; Cánovas et al. 2008, Chávez-Galarza et al. 2017). Haplotypes A16', found in Santa Maria and São Miguel, and A20, found in Flores and São Jorge, were described for the first time by Franck et al. (2001), and have been reported in the northern part of the Atlantic coast of Iberia (Miguel et al. 2007; Chávez-Galarza et al. 2017).

The identification of the haplotypes is based on previously reported band patterns obtained with the *Dra*I test, or with the sequences available on Genbank. In the Macaronesia, the haplotype A9 was firstly reported for the Canary Islands (De La Rúa et al. 1998). Previous studies in the Canary Islands identified the haplotype A9 following the band patterns obtained with the *Dra*I test and due to the similarity with the

fragments reported for haplotype A9 (De La Rúa et al. 1998; 2001). Haplotype A9, obtained with *Dra*I test, was reported in 2006 and later in 2013 for São Miguel the only island of the Azores surveyed (De La Rúa et al. 2006, Muñoz et al. 2013). In 2006, haplotype A30, which presents a band pattern similar to that of A9, was identified in colonies of Brazil (Collet et al. 2006). Although both haplotypes belong to the African lineage, A9 has the P_0 element, typical of sub-lineage A_{II}, while haplotype A30 has the P_1 element, typical of sub-lineage A_{III}. The band patterns of A9 and A30 are similar being the larger restricted fragments hardly distinguishable in a 4% wide-range gel (783 bp in A9 and 768 bp in A30). Given that in previous studies carried out in São Miguel only the *Dra*I test was performed (De la Rúa et al. 2006; Muñoz et al. 2013), it is possible that A9 was misidentified and confounded with A30. Nevertheless, all the remaining haplotypes identified here as A30 (12) will be sequenced to further confirm this haplotype. In the Iberian Peninsula, haplotype A30 has only been reported in northern Portugal (Chávez-Galarza et al. 2017). The presence of this haplotype in Brazil suggests that it was probably introduced by Portuguese settlers (Collet et al. 2006). Haplotype A30, which was private to São Miguel, was likely introduced in the Azores from honey bee populations of northern Portugal.

The African sub-lineage A_I is reported for the first time in the Azores with detection of haplotype A1 solely in Faial. A1 was the first African haplotype to be described (Garnery et al. 1993). This haplotype is widely distributed in sub-Saharan Africa decreasing progressively from Guinea towards south-eastern Africa (Garnery et al. 1993 1995; Franck et al. 1998, 2001) and it is also found in populations of *A. m. sicula* from Sicily (Franck et al. 2000b). This haplotype has been reported both in Spain and Portugal (Garney et al. 1993, 1995; Franck et al. 1998; De La Rúa et al. 1999; Cánovas et al. 2008; Chávez-Galarza et al. 2015; 2017), but it is in Alentejo and Algarve where it reaches the highest frequencies in Iberia (Pinto et al. 2013; Chávez-Galarza et al. 2017). In the Balearic Islands, A1 is fixed in Majorca and it shows high frequencies in Minorca (De La Rúa et al. 2001b). In the Canaries, haplotype A1 occurs across all islands, except in El Hierro (De La Rúa et al. 1998, 2001, 2002; Muñoz et al. 2013), and in a more recent study, it was detected in all localities sampled across La Palma (Miguel et al. 2016). Surprisingly, unlike in the Azores, haplotypes from sub-lineage A_I (A1, A2, A4') are represented at high frequencies in Madeira (De La Rúa et al. 2006; Muñoz et al. 2013), and are also frequent in the Canary Islands (A1, A3) (De La Rúa et al. 1998, 2001, 2002; Muñoz et al. 2013). In fact, A1 is the most common haplotype in Madeira (De La Rúa et al. 2006; Muñoz et al. 2013) and was also found in honey bee populations from Brazil (Collet et al. 2006).

Finally, sub-lineage A_{II} was only represented by haplotypes A8, private to Santa Maria, and A10', private to São Miguel, although the former was earlier reported in São Miguel at low frequency (Muñoz et al. 2013). In the Canary Islands, haplotype A8 was only found in Tenerife, also at low frequencies (Muñoz et al. 2013). This haplotype occurs in the Iberian Peninsula, both in Portugal and Spain (Cánovas et al. 2008; Pinto et al. 2013; Chávez-Galarza et al. 2017), but it is in populations of *A. m. intermissa* of north Africa where it is more frequent (Garnery et al. 1995; Franck et al. 2001; Shaibi et al. 2009).

The least frequent non-African haplotypes in the Azores were those of Western European ancestry (lineage M). Only two different M haplotypes were observed: M70 in São Miguel and M7 in Graciosa and Terceira. Haplotype M70 was recently described by Szalanski et al. (2016) and interestingly, both haplotypes (M70 and M7) were recently found at high frequencies in the Hawaii archipelago (Szalanski et al. 2016). Although M7 is characteristic from the West European M lineage, has been surveyed more frequently in *A. m. ligustica* in northwestern Italy (Franck et al. 2000b) and has been reported in honey bee populations from the north of the Iberian Peninsula (Garnery et al. 1993, 1998; Franck et al. 2001; Miguel et al. 2007; Cánovas et al. 2008; Chávez-Galarza et al. 2017). Furthermore, haplotype M7 was detected in low frequencies in Gran Canaria (Muñoz et al. 2013). Haplotype M70, also from the West European M lineage, was only surveyed in Hawaii and in the northeastern Spain (Chávez-Galarza et al. 2017).

São Miguel has been the single Azorean island that was surveyed earlier (De La Rúa et al. 2006; Pinto et al. 2012; Muñoz et al. 2013), which allowed to infer about temporal changes in the mtDNA. Contrasting with the previous studies, the current study included a higher number of colonies (N=99 against N=48, in De La Rúa et al. (2006), and N=48, in Muñoz et al. 2013), which permitted detection of a large number of haplotypes (8), four of them private, representing three different lineages. Haplotypes A14' and M70 had never been detected before in the island. Although in previous researches were detected the haplotype A11 and A8 (Muñoz et al. 2013) there were no evidence of these haplotypes in the present study. Sequencing of a sub-sample from São Miguel revealed additional mitochondrial variation. In addition to the previously reported haplotype C1 (De La Rúa et al. 2006; Muñoz et al. 2013), there is at least one more haplotype of C-lineage ancestry (C2j) in this island. Nonetheless, the frequency of C-lineage haplotypes is lower than that of earlier surveys. For the first time, lineage M was detected in São Miguel and the frequency of the African sub-

lineage A_{III} increased, in part by a possible misidentification of A9 (sub-lineage A_{II}) instead of A30 (sub-lineage A_{III}).

Similarly, to the Canary Islands (De La Rúa et al. 1998; 2001; Muñoz et al. 2013; Miguel et al. 2016), the frequency of C-lineage haplotypes is highly variable across the eight Azorean islands. Despite the ubiquity of foreign haplotypes in the Azores, it is still possible to observe that in Terceira and São Jorge C-lineage colonies were not detected and in Santa Maria there was only one colony carrying a C haplotype. High frequencies of C haplotypes were observed in Pico, Faial and Graciosa. Events of introductions of foreign honey bee subspecies can be inferred from mtDNA variation through the *Dra*I test. In this study, data from the *Dra*I test and sequencing revealed the presence of C-lineage colonies in six islands (Flores, Faial, Pico, Graciosa, São Miguel and Santa Maria), a fact that was not surprising since beekeepers have shown worldwide a preference for *A. m. ligustica, A. m. carnica, and A. m. caucasia*. Although *A. m. caucasia* belongs to the morphological O lineage (Ruttner 1988), the *Dra*I patterns of this subspecies are not differentiated from those obtained for C lineage (Cornuet and Garnery 1991; Garnery et al. 1992).

In the 1980s, the regional government of the Azores supported a breeding program to improve the local honey bee stock. This program was implemented by a French beekeeper, Jean Pierre Lhérété. Hundreds of queens of *A. m. ligustica* and *A. m. caucasia* were purchased to queen breeders in Italy and France and were introduced in Santa Maria and Graciosa, respectively, where the breeding stations of pure stock were located. The daughters of those pure queens were taken to Pico for large-scale hybridization and multiplication and then distributed across the Azores (personal communication of JP Lhérété to Vincent Douarre). In addition to this authorized large scale introductions of foreign honey bee stock, there are anecdotal reports of illegal importations of commercial C-lineage honey bees from Canada into Pico in the late 1990's, which possibly led to introduction of the mite *Varroa destructor* in the Azores. The official detection of this mite in the Azores was in 2000 in Pico, 2001 in Flores, and 2007 in Faial.

Despite the high frequency of C lineage detected in Graciosa, haplotypes of lineages M and A ancestry, more specifically sub-lineage A_{III}, are still observed in this island. Interestingly, in this island there is a considerable proportion of lineage M, namely haplotype M7, when comparing to the remaining islands. Along with C1 and C2, M7 is one of the most reported foreign haplotypes, characteristic of the commonly worldwide imported susbpecies (Franck et al. 2000b; Muñoz et al. 2013, 2014 a, b).

Sequencing data indicates that within lineage C there are at least three different haplotypes in the Azores: C1, C2 (variant j) and C3 (variant b). Haplotype C1, which is characteristic of the Italian honey bee *A. m. ligustica* (Franck et al. 2000b), was also observed at high frequencies in Tenerife and El Hierro (Canaries) and was previously reported in São Miguel (Muñoz et al. 2013). It is also common in the United States (Delaney et al. 2009; Magnus et al. 2011), Brazil, and Uruguay (Collet et al. 2006). Haplotype C2j was also detected in La Palma (Canaries) and although haplotypes of C-lineage ancestry are very rare in the Iberian Peninsula, a single colony harboring a C2j was found in northern Portugal (Cánovas et al. 2008, Chávez-Galarza et al. 2017). The C3b haplotype, which was first described by Magnus et al. (2011) in the USA, was only detected in two islands of Azores, Faial and Pico, the islands with the highest percentage of C lineage.

A recurrent problem of the introduction of foreign honey bees besides the introgressive hybridization lies in the fact that this practice exposes local population to new pathogens and parasites threatening their survivability (Matheson et al. 1996; Moritz et al. 2005; Munoz et al. 2009; Mutinelli 2011). Colony diseases are one of the most important impediments to the beekeeping activity across the world (Muñoz et al. 2014b). Probably, introductions of foreign honey bees were also responsible for the appearance of V. *destructor* in Pico in 2000 and subsequent dissemination to Flores in 2001 and Faial in 2008. Varroa represents one of the most important threats to apiculture, being associated with colony losses in Europe, but also worldwide (De La Rúa et al. 2009; Rosenkranz et al. 2010; Martin et al. 2012; Francis et al. 2013; Emsen et al. 2015; Meixner et al. 2015; Mordecai et al. 2016). Recently, Segura (2016) examined the 473 colonies used in this study to determine the prevalence of the main pathogenic agents in the Azorean honey bee populations and the relation of Deformed Wing Virus (DWV) with varroa. This author confirmed that varroa is still restricted to Faial, Pico and Flores and also verified its association with DWV.

The introduction of foreign subspecies, followed by introgressive hybridization events, has changed the gene pool of native honey bee populations in many parts of their natural distribution, but especially in Europe (Garnery et al. 1998; De La Rúa et al. 2009, Pinto et al. 2014). The influence of the beekeeping activity in the genetic composition of the Azorean honey bee populations is notorious by the high levels of foreign haplotypes, mainly from C lineage. It would be interesting to establish a program to preserve the gene pool of local honey bees in the Azores (especially in the islands with negligible levels of C-lineage influence), similar to what has been done in

La Palma since 2001, when a conservation program was implemented to preserve the local black honey bee.

4.2. Geometric Morphometrics

It is currently known that the wings venations of the honey bees play a crucial role during the flight, especially the edge veins as they contribute to the stability of the wing (Combes and Daniel 2003). The wings venation was examined in the 473 colonies of the Azores. The results of the wings venation obtained with the powerful MorphoJ, show that colonies from Graciosa always tend to exhibit a marked variation pattern that stands out from the other islands, mainly resembling *A. m. ligustica*, but also *A. m. carnica*. However, it was *A. m. caucasia* queens, which belong to the morphological O lineage (Ruttner 1988), that were introduced in Graciosa in the 1980s (personal communication of JP Lhérété to Vincent Douarre). While the *Dra*I test has identified C-haplotypes in *A. m. caucasia* colonies (Cornuet and Garnery 1991; Garnery et al. 1992), Ruttner (1988) placed *A. m. caucasia* in the Middle Eastern (O) lineage, based on morphological data. Unfortunately, we did not have access to samples of *A. m. caucasia*, which would help understanding the morphological relationships between the honey bees from Graciosa and *A. m. caucasia*.

Although the mitochondrial data showed a high frequency of C-derived haplotypes, especially in Pico, Faial and Graciosa, the wing geometric morphometrics data do not clearly group the colonies from these islands with the C-lineage references *A. m. ligustica* and *A. m. carnica*. This may be due to the fact that wing venations patterns present a biparental inheritance, contrary to the mtDNA (Bonatti et al. 2014). Moreover, it appears that queens from foreign subspecies show a preference for native drones, which are better adapted to the local environmental conditions (Muñoz et al. 2014a).

A great similarity is observed between colonies of Santa Maria and São Miguel which exhibit patterns of shape variation that are to some extent, different from the other islands, showing a closer relationship with the reference *A. m. iberiensis*. Historical reports indicate that the Azores were discovered by Portuguese navigators in the fifteenth century (Matos 1989; Mendonça 1996). The colonization of the archipelago was first initiated in 1439 in Santa Maria and São Miguel with the first settlers coming mainly from Portugal and Madeira (Matos 1989; Mendonça 1996). The establishment of humans in the islands required the introduction of other organisms necessary for human consumption, such as plants, small animals, grazing animals, and honey bees. Honey bees were amongst the most important organisms due to the

products derived from their activity, mainly wax, used to produce candles, and honey. Therefore, Portuguese settlers brought into Santa Maria and São Miguel honey bee colonies most likely originating from northern Portugal, as suggested both by geometric morphometrics and mtDNA data.

Contrary to São Miguel and Santa Maria, differences with populations of *A. m. iberiensis* are more accentuated when compared with Faial, Pico, São Jorge and Flores. Although these populations are more differentiated from *A. m. iberiensis*, they exhibited morphometric patterns similar to Santa Maria and São Miguel, which may *suggest* colonization from these islands instead of from the Iberian Peninsula. The lowest values for Mahalanobis and Procrustes distances were obtained for Faial and Pico and the stronger affinity between these islands is highlighted in the graphical representations. The strong morphological similarity could be explained by the geographical proximity, since these islands are separated by little more than 6 km facilitating gene flow between both populations by natural or human-mediated means.

The analyses of the shape patterns using a new variable including representing the evolutionary lineages and sub-lineages observed for population of the Azores, *A. m. iberiensis*, *A. m. ligustica* and *A. m. carnica* clearly shows relationship of the population of the Azores with the population from *A. m. iberiensis*. However, the high levels of mtDNA introgression are also reflected in the wing shape patterns since these islands overlapped with the subspecies *A. m. ligustica* and *A. m. carnica*, both C-lineage.

Finally, although the honey bee populations of the Azores show a close relationship with *A. m. iberiensis,* they present variations in the wing venation patterns that clearly distinguish them from the populations in the mainland. Several non-mutually exclusive factors can account for this pattern: the founder effect resulting from the colonization process and the recent introduction of foreign subspecies as *A. m. ligustica* and *A. m. carnica* is possibly contributing to the differences observed. Furthermore, environmental factors such as climatic conditions, mainly temperature and humidity, can potentially shape the patterns of wing morphology (Debat et al. 2003; Pitchers et al. 2013; Bonatti et al. 2014; Perrard et al. 2014; Liu et al. 2016). Moreover, the geographic isolation of the Azores has considerably limited gene flow, which may have also contributed to the differences in the wing shape among the continental and insular populations, similar to what was observed for other insect species (Camara et al. 2006; Perrard et al. 2014). Despite the lack of suitable wing geometric morphometric studies in honey bees from the Iberian Peninsula, studies have demonstrated a great power of discrimination within different populations of *A. m. iberiensis* in the Iberian

Peninsula (Chávez-Galarza et al. 2016b) being congruent with the molecular analysis (Chávez-Galarza et al. 2013, 2015). Also results from geometric morphometrics data corroborate the molecular data in a study performed in honey bees from La Palma, Canary (Miguel et al. 2016).

In general, both molecular and morphometric data were able to detect structure between the eight islands of Azores. Similar to the honey bee populations from Madeira and the Canary Islands (De La Rúa et al. 1998, 2006; Muñoz et al. 2013; Miguel et al. 2016), in the Azores, the maternal sub-lineage A_{III} is the most spread. Haplotypes of sub-lineage A_{III} ancestry are absent in Eastern Europe and Western Europe, north of the Pyrenees, and are rare in Africa (Franck et al. 2001; De la Rúa et al. 2006) and in most of the Iberian Peninsula (Cánovas et al. 2008; Miguel et al. 2007). However, a high frequency of A_{III} haplotypes is found in the Atlantic side of Iberia north of Lisbon (Pinto et al. 2013; Chávez-Galarza et al. 2017), making Portugal the largest reservoir of these haplotypes. Also, the geometric morphometrics data reveals a closer relationship of the populations from the Azores with those from continental Portugal than with those from Spain. These results support the historical records reporting the introduction of honey bee colonies by Portuguese settlers in the 15^{th} century and suggest that colonies, carrying haplotypes from sub-lineage A_{III} were brought from northern Portugal. The presence of haplotype from C lineage, mainly due to the importation of *A. m. ligustica* and *A. m. carnica,* are not suitable to infer a historical relationship, since these haplotypes were recently introduced in the islands and are widely distributed across the world. Moreover, haplotypes from C lineage are very rare or even non-existent populations in honey bee populations from Iberian Peninsula (Cánovas et al. 2008, Pinto et al. 2012, 2013; Chávez-Galarza et al. 2017), the putative source of the honey bee from Azores archipelago.

Despite the colonization of the islands from continental Portugal, the honey bees from the Azores present differences in molecular and morphological traits. Founder effects that occurred during the process of colonization of the different islands may explain the differences observed between these populations and populations from the mainland. It is also possible that adaptation to the insular conditions may have led to divergence between mainland and insular populations.

5. Conclusion

This study was the first to survey the eight Azorean islands populated by honey bees, allowing to infer maternal genetic diversity and the wing geometric morphometric patterns of the *A. mellifera* in this Macaronesian archipelago. The results reveal a consistency between morphological and genetic data and suggest that the honey bees of the Azores were likely introduced from northern Portugal.

The mitochondrial intergenic $tRNA^{leu}$-cox2 region assessed with the *Dra*I test revealed to be highly informative to infer the genetic variability of the honey bee populations from the Azores. Although haplotypes from the African sub-lineage A_{III} are the most frequently observed in the Azores, the contemporary influence of beekeeping is mainly reflected by the high percentage of colonies carrying mitochondria of eastern European (C-lineage) ancestry in many islands.

Geometric morphometrics techniques also demonstrated to be a suitable and efficient tool to study wing shape differences, confirming the efficiency of the wing venation to identify variations among populations from the different islands and mainland. Although the morphometric data suggest that historical colonization of the Azores occurred from populations of *A. m. iberiensis*, mainly from Portugal, honey bees from the Azores have differentiated from those of the mainland.

A founder effect, occurring in the historical introduction of honey bee colonies, along with the barrier to gene flow between Azorean and Iberian honey bee populations and a possible adaptation to the insular conditions may underlie differentiation of the Azorean honey bees. Similar findings were obtained by Miguel et al. (2016) with wing geometric morphometrics, microsatellites, and the mtDNA intergenic region in La Palma (Canary Islands). The authors also found a close relationship between honey bee populations of La Palma and Portugal.

Future studies of wing morphometrics geometrics of the Azores should use additional reference subspecies (especially *A. m. caucasia*) and larger samples sizes to better infer the impact of C-lineage honey bees and to elucidate about the presence or absence of the morphological lineage O.

Further investigation using nuclear markers such as microsatellites or SNPs are required to improve current knowledge about the history and demographics of the Azorean honey bees, and about possible genes under selection related with adaptation to the particular insular environment.

A conservation program can help protecting the locally adapted honey bees from the Azores, by reducing or eliminating the colonies of C-lineage ancestry, contributing to a sustainable beekeeping activity.

6. References

Alburaki, M., Moulin, S., Legout, H., Alburaki, A., & Garnery, L. (2011). Mitochondrial structure of Eastern honey bee populations from Syria, Lebanon and Iraq. *Apidologie*, *42*(5), 628.

Alpatov, W. W. (1929). Biometrical studies on variation and races of the honey bee (*Apis mellifera* L.). *The Quarterly Review of Biology*, *4*(1), 1-58.

Anstey, R. L., & Pachut, J. F. (2004). Cladistic and phenetic recognition of species in the Ordovician bryozoan genus Peronopora. *Journal of Paleontology*, *78*(4), 651-674.

Arias, M. C., & Sheppard, W. S. (1996). Molecular Phylogenetics of Honey Bee Subspecies (*Apis mellifera* L.) Inferred from Mitochondrial DNA Sequence. *Molecular phylogenetics and evolution*, *5*(3), 557-566.

Arias, M. C., & Sheppard, W. S. (2005). Phylogenetic relationships of honey bees (Hymenoptera: Apidae: Apini) inferred from nuclear and mitochondrial DNA sequence data. *Molecular phylogenetics and evolution*, *37*(1), 25-35.

Atchley, W. R., & Hall, B. K. (1991). A model for development and evolution of complex morphological structures. *Biological Reviews*, *66*(2), 101-157.

Aytekin, M. A., Terzo, M., Rasmont, P., & Çağatay, N. (2007). Landmark based geometric morphometric analysis of wing shape in Sibirico bombus Vogt (Hymenoptera: Apidae: *Bombus Latreille*). In *Annales de la Société Entomologique de France* (Vol. 43, No. 1, pp. 95-102). Taylor & Francis Group.

Baer, B. (2005). Sexual selection in Apis bees. *Apidologie*, *36*(2), 187-200.

Barchuk, A. R., Cristino, A. S., Kucharski, R., Costa, L. F., Simões, Z. L., & Maleszka, R. (2007). Molecular determinants of caste differentiation in the highly eusocial honey bee Apis mellifera. *BMC developmental biology*, *7*(1), 70.

Barour, C., & Baylac, M. (2016). Geometric morphometric discrimination of the three African honey bee subspecies *Apis mellifera intermissa*, *A. m. sahariensis* and *A. m. capensis* (Hymenoptera, Apidae): Fore wing and hind wing landmark configurations. *Journal of Hymenoptera Research*, *52*, 61.

Barton, N. H., & Mallet, J. (1996). Natural selection and random genetic drift as causes of evolution on islands [and discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *351*(1341), 785-795.

Baylac, M., Villemant, C., & Simbolotti, G. (2003). Combining geometric morphometrics with pattern recognition for the investigation of species complexes. *Biological Journal of the Linnean Society*, *80*(1), 89-98.

Boc, A., & Makarenkov, V. (2012). T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic acids research*, *40*(W1), W573-W579.

Boecking, O., & Genersch, E. (2008). Varroosis–the ongoing crisis in bee keeping. *Journal für Verbraucherschutz und Lebensmittelsicherheit*, *3*(2), 221-228.

Bonatti, V., Simões, Z. L. P., Franco, F. F., & Francoy, T. M. (2014). Evidence of at least two evolutionary lineages in *Melipona subnitida* (Apidae, Meliponini) suggested by mtDNA variability and geometric morphometrics of forewings. *Naturwissenschaften*, *101*(1), 17-24.

Bookstein, F. L. (1991). Morphometric tools for landmark data Cambridge University Press. *New York*.

Bookstein, F. L. (1996). Combining the tools of geometric morphometrics. In *Advances in morphometrics* (pp. 131-151). Springer US.

Bookstein, F. L., Sampson, P. D., Streissguth, A. P., & Connor, P. D. (2001). Geometric morphometrics of corpus callosum and subcortical structures in the fetal-alcohol-affected brain. *Teratology*, *64*(1), 4-32.

Boomsma, J. J., Baer, B., & Heinze, J. (2005). The evolution of male traits in social insects. *Annual Review Entomology*, *50*, 395-420.

Bouga, M., Kilias, G., Harizanis, P. C., Papasotiropoulos, V., & Alahiotis, S. (2005a). Allozyme variability and phylogenetic relationships in honey bee (Hymenoptera: Apidae: *Apis mellifera*) populations from Greece and Cyprus. *Biochemical Genetics*, *43*(9), 471-483.

Bouga, M., Harizanis, P. C., Kilias, G., & Alahiotis, S. (2005b). Genetic divergence and phylogenetic relationships of honey bee *Apis mellifera* (Hymenoptera: Apidae) populations from Greece and Cyprus using PCR-RFLP analysis of three mtDNA segments. *Apidologie*, *36*(3), 335-344.

Breeze, T. D., Bailey, A. P., Balcombe, K. G., & Potts, S. G. (2011). Pollination services in the UK: How important are honey bees?. *Agriculture, Ecosystems & Environment*, *142*(3), 137-143.

Bruner, E. (2004). Geometric morphometrics and paleoneurology: brain shape evolution in the genus Homo. *Journal of Human Evolution*, *47*(5), 279-303.

Byatt, M. A., Chapman, N. C., Latty, T., & Oldroyd, B. P. (2016). The genetic consequences of the anthropogenic movement of social bees. *Insectes Sociaux*, *63*(1), 15-24.

Calado, H., Borges, P., Phillips, M., Ng, K., & Alves, F. (2011). The Azores archipelago, Portugal: improved understanding of small island coastal hazards and mitigation measures. *Natural Hazards*, *58*(1), 427-444.

Camara, M., Caro-Riano, H., Ravel, S., Dujardin, J. P., Hervouet, J. P., Meeus, T. D., Kagbadouno, M. S., Bouyer, J., & Solano, P. (2006). Genetic and Morphometric Evidence for Population Isolation of *Glossina palpalis gambiensis* (Diptera: Glossinidae) on the Loos Islands, Guinea. *Journal of Medical Entomology*, *43*(5), 853-860.

Cánovas, F., De la Rúa, P., Serrano, J., & Galián, J. (2008). Geographical patterns of mitochondrial DNA variation in *Apis mellifera iberiensis* (Hymenoptera: Apidae). *Journal of Zoological Systematics and Evolutionary Research*, *46*(1), 24-30.

Cánovas, F., de la Rúa, P., Serrano, J., & Galián, J. (2011). Microsatellite variability reveals beekeeping influences on Iberian honey bee populations. *Apidologie*, *42*(3), 235-251.

Chávez-Galarza, J., Henriques, D., Johnston, J. S., Azevedo, J. C., Patton, J. C., Muñoz, I., De la Rúa, P. & Pinto, M. A. (2013). Signatures of selection in the Iberian honey bee (*Apis mellifera iberiensis*) revealed by a genome scan analysis of single nucleotide polymorphisms. *Molecular Ecology*, *22*(23), 5890-5907.

Chávez-Galarza, J., Henriques, D., Johnston, J. S., Carneiro, M., Rufino, J., Patton, J. C., & Pinto, M. A. (2015). Revisiting the Iberian honey bee (Apis mellifera iberiensis) contact zone: maternal and genome-wide nuclear variations provide support for secondary contact from historical refugia. *Molecular ecology*, *24*(12), 2973-2992.

Chávez Galarza, J. C. (2016a). Population genomics and landscape genetics of the Iberian honey bee (*Apis mellifera iberiensis*).

Chávez-Galarza, J., Galaschi-Teixeira, J. S., Neves, C., Henriques, D., Francoy, T. M. & Pinto, M. A. (2016b). Geometric morphometrics of wings of drones and workers support a process of secondary contact in the Iberian honey bee (*Apis mellifera iberiensis*). EURBEE 2016: 7th European Conference of Apidology. 7-9 Sept. 2016, Cluj-Napoca, Romania.

Chávez-Galarza, J., Garnery, L., Henriques, D., Neves, C. J., Loucif-Ayad, W., Jonhston, J. S., & Pinto, M. A. (2017). Mitochondrial DNA variation of *Apis mellifera iberiensis*: further insights from a large-scale study using sequence data of the tRNAleu-cox2 intergenic region. *Apidologie*, 1-12.

Chen, C., Liu, Z., Pan, Q., Chen, X., Wang, H., Guo, H., Liu, S., Lu, H., Tian, S., Li, R. & Shi, W. (2016). Genomic analyses reveal demographic history and temperate

adaptation of the newly discovered honey bee subspecies *Apis mellifera sinisxinyuan* n. ssp. *Molecular biology and evolution*, *33*(5), 1337-1348.

Clabaut, C., Bunje, P. M., Salzburger, W., & Meyer, A. (2007). Geometric morphometric analyses provide evidence for the adaptive character of the Tanganyikan cichlid fish radiations. *Evolution*, *61*(3), 560-578.

Claude, J., Pritchard, P. C., Tong, H., Paradis, E., & Auffray, J. C. (2004). Ecological correlates and evolutionary divergence in the skull of turtles: a geometric morphometric assessment. *Systematic Biology*, *53*(6), 933-948.

Collet, T., Ferreira, K. M., Arias, M. C., Soares, A. E. E., & Del Lama, M. A. (2006). Genetic structure of Africanized honeybee populations (*Apis mellifera* L.) from Brazil and Uruguay viewed through mitochondrial DNA COI-COII patterns. *Heredity*, *97*(5), 329.

Combes, S. A., & Daniel, T. L. (2003). Flexural stiffness in insect wings I. Scaling and the influence of wing venation. *Journal of experimental biology*, *206*(17), 2979-2987.

Cornuet, J. M., & Fresnaye, J. (1989). Biometrical study of honey bee populations from Spain and Portugal. *Apidologie*, 20 (1), 393-403.

Cornuet, J. M., & Garnery, L. (1991). Mitochondrial DNA variability in honey bees and its phylogeographic implications. *Apidologie*, *22*(6), 627-642.

Cridland, J. M., Tsutsui, N. D., & Ramírez, S. R. (2017). The Complex Demographic History and Evolutionary Origin of the Western Honey Bee, *Apis Mellifera*. *Genome Biology and Evolution*, *9*(2), 457-472.

Crozier, R. H., & Crozier, Y. C. (1993). The mitochondrial genome of the honey bee *Apis mellifera*: complete sequence and genome organization. *Genetics*, *133*(1), 97-117.

Crozier, Y. C., Koulianos, S., & Crozier, R. H. (1991). An improved test for Africanized honeybee mitochondrial DNA. *Cellular and Molecular Life Sciences*, *47*(9), 968-969.

Debat, V., Bégin, M., Legout, H., & David, J. R. (2003). Allometric and nonallometric components of *Drosophila* wing shape respond differently to developmental temperature. *Evolution*, *57*(12), 2773-2784.

De la Rúa, P., Serrano, J., & Galián, J. (1998). Mitochondrial DNA variability in the Canary Islands honey bees (*Apis mellifera* L.). *Molecular ecology*, *7*(11), 1543-1547.

De la Rúa P., Galián, J. & Serrano, J. (1999) Variabilidad del ADN mitocondrial en poblaciones de abejas de la miel *(Apis mellifera* L.) de la Región de Murcia. *Investigación agraria. Producción y sanidad animales*, *9*, 43-49.

De la Rúa, P., Galián, J., Serrano, J., & Moritz, R. F. A. (2001a). Genetic structure and distinctness of *Apis mellifera* L. populations from the Canary Islands. *Molecular Ecology*, *10*(7), 1733-1742.

De La Rúa, P., Galián, J., Serrano, J., & Moritz, R. F. A. (2001b). Molecular characterization and population structure of the honeybees from the Balearic islands (Spain). *Apidologie*, *32*(5), 417-427.

De la Rúa, P., Galián, J., Serrano, J., & Moritz, R. F. A. (2003). Genetic structure of Balearic honey bee populations based on microsatellite polymorphism. *Genetics Selection Evolution*, *35*(3), 339.

De La Rúa, P., Jiménez, Y., Galián, J., & Serrano, J. (2004). Evaluation of the biodiversity of honey bee (*Apis mellifera*) populations from eastern Spain. *Journal of apicultural research*, *43*(4), 162-166.

De La Rúa, P., Hernández-García, R., Jiménez, Y., Galián, J., & Serrano, J. (2005). Biodiversity of *Apis mellifera iberica* (Hymenoptera: Apidae) from northeastern Spain assessed by mitochondrial analysis. *Insect Systematics & Evolution*, *36*(1), 21-28.

De La Rúa, P., Galián, J., Pedersen, B. V., & Serrano, J. (2006). Molecular characterization and population structure of *Apis mellifera* from Madeira and the Azores. *Apidologie*, *37*(6), 699.

De la Rúa, P., Radloff, S., Hepburn, R., & Serrano, J. (2007). Do molecular markers support morphometric and pheromone analyses? A preliminary case study in Apis mellifera populations of Morocco. *Archivos de zootecnia*, *56*(213), 1-11.

De la Rúa, P., Jaffé, R., Dall'Olio, R., Muñoz, I., & Serrano, J. (2009). Biodiversity, conservation and current threats to European honey bees. *Apidologie*, *40*(3), 263-284.

Delaney, D. A., Meixner, M. D., Schiff, N. M., & Sheppard, W. S. (2009). Genetic characterization of commercial honey bee (Hymenoptera: Apidae) populations in the United States by using mitochondrial and microsatellite markers. *Annals of the Entomological Society of America*, *102*(4), 666-673.

De Meulemeester, T., Michez, D., Aytekin, A. M., & Danforth, B. N. (2012). Taxonomic affinity of halictid bee fossils (Hymenoptera: Anthophila) based on geometric morphometrics analyses of wing shape. *Journal of Systematic Palaeontology*, *10*(4), 755-764.

DuPraw, E. J. (1964). Non-linnean taxonomy. *Nature*, *202*, 849-852.

DuPraw, E. J. (1965). Non-Linnean taxonomy and the systematics of honey bees. *Systematic Zoology*, *14*(1), 1-24.

Emerson, B. C. (2002). Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular ecology*, *11*(6), 951-966.

Emsen, B., Hamiduzzaman, M. M., Goodwin, P. H., & Guzman-Novoa, E. (2015). Lower virus infections in Varroa destructor-infested and uninfested brood and adult honey bees (*Apis mellifera*) of a low mite population growth colony compared to a high mite population growth colony. *Plos one*, *10*(2),1-10.

Engel, M. S. (1999). The taxonomy of recent and fossil honey bees (Hymenoptera: Apidae; Apis).

Erlandsson, S. (1979). *Bombus canariensis* Pérez, 1895 n. stat and *B. maderensis* n. sp. from the Macaronesian Islands. *Insect Systematics & Evolution*, *10*(3), 187-192.

Fernández-Palacios, J. M., de Nascimento, L., Otto, R., Delgado, J. D., García-del-Rey, E., Arévalo, J. R., & Whittaker, R. J. (2011). A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography*, *38*(2), 226-246.

Figueiredo, B., Palmqvist, P., & Pérez-Claros, J. A. (2009). Ecomorphological correlates of craniodental variation in bears and paleobiological implications for extinct taxa: an approach based on geometric morphometrics. *Journal of Zoology*, *277*(1), 70-80.

Foster, D. J., Podos, J., & Hendry, A. P. (2008). A geometric morphometric appraisal of beak shape in Darwin's finches. *Journal of evolutionary biology*, *21*(1), 263-275.

Francis, R. M., Nielsen, S. L., & Kryger, P. (2013). Varroa-virus interaction in collapsing honey bee colonies. *PloS one*, *8*(3), e57540.

Franck, P., Garnery, L., Solignac, M., and J. M. Cornuet (2000a). Molecular confirmation of a fourth lineage in honey bees from the Middle East. *Apidologie*, *31*, 167-180.

Franck, P., Garnery, L., Celebrano, G., Solignac, M., & Cornuet, J. M. (2000b). Hybrid origins of honey bees from Italy (*Apis mellifera ligustica*) and Sicily (*A. m. sicula*). *Molecular Ecology*, *9*(7), 907-921.

Franck, P., Garnery, L., Loiseau, A., Oldroyd, B. P., Hepburn, H. R., Solignac, M., & Cornuet, J. M. (2001). Genetic diversity of the honey bee in Africa: microsatellite and mitochondrial data. *Heredity*, *86*(4), 420-430.

Francoy, T. M., Wittmann, D., Drauschke, M., Müller, S., Steinhage, V., Bezerra-Laure, M. A., De Jong, D. & Gonçalves, L. S. (2008). Identification of Africanized honey bees through wing morphometrics: two fast and efficient procedures. *Apidologie*, *39*(5), 488-494.

Francoy, T. M., Wittmann, D., Steinhage, V., Drauschke, M., Müller, S., Cunha, D. R., Nascimento, A. M., Figueiredo V.L.C., Simões, Z.L.P., De Jong, D., Arias, M. C. & Gonçalves, L.S. (2009). Morphometric and genetic changes in a population of

Apis mellifera after 34 years of Africanization. *Genetics and Molecular Research*, *8*(2), 709-717.

Garnery, L., Cornuet, J. M., & Solignac, M. (1992). Evolutionary history of the honey bee *Apis mellifera* inferred from mitochondrial DNA analysis. *Molecular ecology*, *1*(3), 145-154.

Garnery, L., Solignac, M., Celebrano, G., & Cornuet, J. M. (1993). A simple test using restricted PCR-amplified mitochondrial DNA to study the genetic structure of *Apis mellifera* L. *Cellular and Molecular Life Sciences*, *49*(11), 1016-1021.

Garnery, L., Mosshine, E. H., Oldroyd, B. P., & Cornuet, J. M. (1995). Mitochondrial DNA variation in Moroccan and Spanish honey bee populations. *Molecular Ecology*, *4*(4), 465-472.

Garnery, L., Franck, P., Baudry, E., Vautrin, D., Cornuet, J. M., & Solignac, M. (1998). Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*) I. Mitochondrial DNA. *Genetics Selection Evolution*, *30*, 31-47.

Gerard, M., Michez, D., Fournier, D., Maebe, K., Smagghe, G., Biesmeijer, J. C., & De Meulemeester, T. (2015). Discrimination of haploid and diploid males of *Bombus terrestris* (Hymenoptera; Apidae) based on wing shape. *Apidologie*, *46*(5), 644-653.

Ginter, C. C., DeWitt, T. J., Fish, F. E., & Marshall, C. D. (2012). Fused traditional and geometric morphometrics demonstrate pinniped whisker diversity. *PloS one*, *7*(4), e34481.

Goetze, G. (1940). *Die beste Biene: Zuechtungs-und Rassen-kunde der Honigbiene nach dem heutigen Stand von Wissenschaft und Praxis*. Liedloff, Loth und Michaelis.

Gomez-Robles, A., Martinon-Torres, M., de Castro, J. M. B., Prado, L., Sarmiento, S., & Arsuaga, J. L. (2008). Geometric morphometric analysis of the crown morphology of the lower first premolar of hominins, with special attention to Pleistocene Homo. *Journal of human evolution*, *55*(4), 627-638.

Gonzalez, P. N., Bernal, V., & Perez, S. I. (2009). Geometric morphometric approach to sex estimation of human pelvis. *Forensic Science International*, *189*(1), 68-74.

Grozinger, C. M., Fan, Y., Hoover, S. E., & Winston, M. L. (2007). Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Molecular Ecology*, *16*(22), 4837-4848.

Gumiel, M., Catala, S., Noireau, F., Rojas de Arias, A., Garcia, A., & Dujardin, J. P. (2003). Wing geometry in *Triatoma infestans* (Klug) and *T. melanosoma* Martinez, Olmedo & Carcavallo (Hemiptera: Reduviidae). *Systematic Entomology*, *28*(2), 173-180.

Hall, H. G., & Smith, D. R. (1991). Distinguishing African and European honey bee matrilines using amplified mitochondrial DNA. *Proceedings of the National Academy of Sciences*, *88*(10), 4548-4552.

Han, F., Wallberg, A., & Webster, M. T. (2012). From where did the Western honey bee (*Apis mellifera*) originate?. *Ecology and evolution*, *2*(8), 1949-1957.

Harpur, B. A., Kent, C. F., Molodtsova, D., Lebon, J. M., Alqarni, A. S., Owayss, A. A., & Zayed, A. (2014). Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proceedings of the National Academy of Sciences*, *111*(7), 2614-2619.

Harvati, K. (2004). 3-D geometric morphometric analysis of temporal bone landmarks in Neanderthals and modern humans. In *Morphometrics* (pp. 245-258). Springer Berlin Heidelberg.

Hepburn, H. R., & Radloff, S. E. (1998). *Honey bees of Africa*. Springer; Berlin, Germany. 370 pp.

Hoffmann, A. A., & Shirriffs, J. (2002). Geographic variation for wing shape in *Drosophila serrata*. *Evolution*, *56*(5), 1068-1073.

Ivanova, E. N., Staykova, T. A., & Bouga, M. (2007). Allozyme variability in honey bee populations from some mountainous regions in the southwest of Bulgaria. *Journal of apicultural research*, *46*(1), 3-7.

Ivanova, E. N., Petrov, P., Bouga, M., Emmanouel, N. G., Tunca, R. I., & Kence, M. (2010). Genetic variation in honey bee (*Apis mellifera* L.) populations from Bulgaria. *Journal of Apicultural Science*, *54*(2), 49-60.

Jaffé, R., Dietemann, V., Allsopp, M. H., Costa, C., Crewe, R. M., Dall'Olio, R., De la Rúa, P., El-Niweiri, M.A.A., Fries, I., Kezic, N., Meusel, M. S., Paxton, R.J., Shaibi, T., Stolle, E. & Moritz, R. F. A. (2010). Estimating the density of honey bee colonies across their natural range to fill the gap in pollinator decline censuses. *Conservation biology*, *24*(2), 583-593.

Jensen, A. B., Palmer, K. A., Boomsma, J. J., & Pedersen, B. V. (2005). Varying degrees of *Apis mellifera ligustica* introgression in protected populations of the black honey bee, *Apis mellifera mellifera*, in northwest Europe. *Molecular Ecology*, *14*(1), 93-106.

Juan, C., Ibrahim, K. M., Oromi, P., & Hewitt, G. M. (1996). Mitochondrial DNA sequence variation and phylogeography of Pimelia darkling beetles on the island of Tenerife (Canary Islands). *Heredity*, *77*(6).

Kandemir, İ., Özkan, A., & Fuchs, S. (2011). Reevaluation of honeybee (*Apis mellifera*) microtaxonomy: a geometric morphometric approach. *Apidologie*, *42*(5), 618.

Klein, A. M., Vaissiere, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society of London B: Biological Sciences*, *274*(1608), 303-313.

Klingenberg, C. P., Leamy, L. J., Routman, E. J., & Cheverud, J. M. (2001). Genetic architecture of mandible shape in mice: effects of quantitative trait loci analyzed by geometric morphometrics. *Genetics*, *157*(2), 785-802.

Kraus, F. B., Neumann, P., Van Praagh, J., & Moritz, R. F. A. (2004). Sperm limitation and the evolution of extreme polyandry in honeybees (*Apis mellifera* L.). *Behavioral Ecology and Sociobiology*, *55*(5), 494-501.

Laidlaw, H. H., & Page, R. E. (1984). Polyandry in honey bees (*Apis mellifera* L.): sperm utilization and intracolony genetic relationships. *Genetics*, *108*(4), 985-997.

Liu, M., Ma, N., & Hua, B. Z. (2016). Intraspecific morphological variation of the scorpionfly *Dicerapanorpa magna* (Chou) (Mecoptera: Panorpidae) based on geometric morphometric analysis of wings. *Contributions to Zoology*, *85*(1), 1-11.

Lo, N., Gloag, R. S., Anderson, D. L., & Oldroyd, B. P. (2010). A molecular phylogeny of the genus Apis suggests that the Giant Honey Bee of the Philippines, *A. breviligula* Maa, and the Plains Honey Bee of southern India, *A. indica* Fabricius, are valid species. *Systematic Entomology*, *35*(2), 226-233.

Losey, J. E., & Vaughan, M. (2006). The economic value of ecological services provided by insects. *Bioscience*, *56*(4), 311-323.

Losos, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature*, *457*(7231), 830-836.

Magnus, R. M., Tripodi, A. D., & Szalanski, A. L. (2011). Mitochondrial DNA diversity of honey bees, *Apis mellifera* L.(Hymenoptera: Apidae) from queen breeders in the United States. *Journal of Apicultural Science*, *55*(1), 37-46.

Marcus, L. F., & Corti, M. (1996). Overview of the new, or geometric morphometrics. In *Advances in morphometrics* (pp. 1-13). Springer US.

Martin, S. J., Highfield, A. C., Brettell, L., Villalobos, E. M., Budge, G. E., Powell, M., Nikaido, S. & Schroeder, D. C. (2012). Global honey bee viral landscape altered by a parasitic mite. *Science, 336*, 1304-1306.

Massot, A. (2015). A agricultura do arquipélago dos Açores (Delegação da COMAGRI). *European Parliament*.

Matias, A., De la Riva, J. X., Torrez, M., & Dujardin, J. P. (2001). *Rhodnius robustus* in Bolivia identified by its wings. *Memorias do Instituto Oswaldo Cruz*, *96*(7), 947-950.

Matos, A. T. (1989). Settlement and Colonization of the Azores. *Portugal no Mundo*, *12*, 185-188.

Meixner, M. D., Leta, M. A., Koeniger, N., & Fuchs, S. (2011). The honey bees of Ethiopia represent a new subspecies of *Apis mellifera — Apis mellifera simensis* n. ssp. *Apidologie*, *42*(3), 425-437.

Meixner, M. D., Pinto, M. A., Bouga, M., Kryger, P., Ivanova, E., & Fuchs, S. (2013). Standard methods for characterising subspecies and ecotypes of *Apis mellifera*. *Journal of Apicultural Research*, *52*(4), 1-28.

Meixner, M. D., Kryger, P., & Costa, C. (2015). Effects of genotype, environment, and their interactions on honey bee health in Europe. *Current Opinion in Insect Science*, *10*, 177-184.

Melo, A., Cruz, J. V., & Coutinho, R. (2014). River water composition in volcanic islands: characterization and main hydrogeochemical processes in São Miguel island (Azores, Portugal). *Comunicações Geológicas*, *101*, 693-696.

Mendonça, L. (1996). História dos Açores. Visão geral (sécs. XV–XIX). *Centro de Apoio Tecnológico à Educação, Ponta Delgada.*

Michez, D., De Meulemeester, T., Rasmont, P., Nel, A., & Patiny, S. (2009). New fossil evidence of the early diversification of bees: Paleohabropoda oudardi from the French Paleocene (Hymenoptera, Apidae, Anthophorini). *Zoologica Scripta*, *38*(2), 171-181.

Miguel, I., Iriondo, M., Garnery, L., Sheppard, W. S., & Estonba, A. (2007). Gene flow within the M evolutionary lineage of *Apis mellifera*: role of the Pyrenees, isolation by distance and post-glacial re-colonization routes in the western Europe. *Apidologie*, *38*(2), 141-155.

Miguel, I., Baylac, M., Iriondo, M., Manzano, C., Garnery, L., & Estonba, A. (2010). Both geometric morphometric and microsatellite data consistently support the differentiation of the *Apis mellifera* M evolutionary branch. *Apidologie*, *42*, 150-161.

Miguel, I., Garnery, L., Iriondo, M., Baylac, M., Manzano, C., Steve Sheppard, W., & Estonba, A. (2015). Origin, evolution and conservation of the honey bees from La Palma Island (Canary Islands): molecular and morphological data. *Journal of Apicultural Research*, *54*(5), 427-440.

Miller, M.P. (1997). Tools for population genetic analyses (TFPGA): a windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by author, *4*, 157.

Mordecai, G. J., Brettell, L. E., Martin, S. J., Dixon, D., Jones, I. M., & Schroeder, D. C. (2016). Superinfection exclusion and the long-term survival of honey bees in Varroa-infested colonies. *The ISME journal*, *10*(5), 1-10.

Muñoz, I., Pinto, M. A., & De la Rúa, P. (2013). Temporal changes in mitochondrial diversity highlights contrasting population events in Macaronesian honey bees. *Apidologie*, *44*(3), 295-305.

Muñoz, I., Pinto, M. A., & De la Rúa, P. (2014a). Effects of queen importation on the genetic diversity of Macaronesian island honey bee populations (*Apis mellifera* Linneaus 1758). *Journal of Apicultural Research*, *53*(2), 296-302.

Muñoz, I., Lodesani, M., & De la Rúa, P. (2014b). Estimating introgression in *Apis mellifera siciliana* populations: are the conservation islands really effective?. *Insect conservation and diversity*, *7*(6), 563-571.

Muñoz, I., Cepero, A., Pinto, M. A., Martín-Hernández, R., Higes, M., & De la Rúa, P. (2014c). Presence of *Nosema ceranae* associated with honey bee queen introductions. *Infection, Genetics and Evolution*, *23*, 161-168.

Mutanen, M., & Pretorius, E. (2007). Subjective visual evaluation vs. traditional and geometric morphometrics in species delimitation: a comparison of moth genitalia. *Systematic Entomology*, *32*(2), 371-386.

Nielsen, D. I., Ebert, P. R., Page, R. E., Hunt, G. J., & Guzmán-Novoa, E. (2000). Improved polymerase chain reaction-based mitochondrial genotype assay for identification of the Africanized honey bee (Hymenoptera: Apidae). *Annals of the Entomological Society of America*, *93*(1), 1-6.

O'Higgins, P. A. U. L., & Jones, N. (1998). Facial growth in *Cercocebus torquatus*: an application of three-dimensional geometric morphometric techniques to the study of morphological variation. *Journal of Anatomy*, *193*(2), 251-272.

Oleksa, A., & Tofilski, A. (2015). Wing geometric morphometrics and microsatellite analysis provide similar discrimination of honey bee subspecies. *Apidologie*, *46*(1), 49-60.

Ottoni, C., Flink, G. L., Evin, A., Geörg, C., De Cupere, B., Van Neer, W., Bartosiewicz, L., Linderholm, A., Barnett, R., Peters, J., Decorte, R., Waelkens, M., Vanderheyden, N., Ricaut, F., Çakirlar, C., Çevik, Ö., Hoelzel, A. R., Mashkour, M., Karimlu, A. F. M., Seno, S. S., Daujat, J., Brock, F., Pinhasi, R., Hongo, H., Perez-Enciso, M., Rasmussen, M., Frantz, L., Megens, H., Crooijmans, R., Groenen, M., Arbuckle, B., Benecke, N., Vidarsdottir, U. A., Burger, J., Cucchi, T., Dobney, K. & Larson, G. (2013). Pig domestication and human-mediated dispersal in western Eurasia revealed through ancient DNA and geometric morphometrics. *Molecular biology and evolution*, *30*(4), 824-832.

Palmer, M. R., Smith, D. R., & Kaftanoglu, O. (2000). Brief communication. Turkish honey bees: genetic variation and evidence for a fourth lineage of *Apis mellifera* mtDNA. *Journal of Heredity*, *91*(1), 42-46.

Peakall R., Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*, 2537–2539.

Perez, S. I., Bernal, V., & Gonzalez, P. N. (2006). Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of anatomy*, *208*(6), 769-784.

Perrard, A., Villemant, C., Carpenter, J. M., & Baylac, M. (2012). Differences in caste dimorphism among three hornet species (Hymenoptera: Vespidae): forewing size, shape and allometry. *Journal of evolutionary biology*, *25*(7), 1389-1398.

Perrard, A., Baylac, M., Carpenter, J. M., & Villemant, C. (2014). Evolution of wing shape in hornets: why is the wing venation efficient for species identification?. *Journal of evolutionary biology*, *27*(12), 2665-2675.

Pinto, M. A., Rubink, W. L., Coulson, R. N., Patton, J. C., & Johnston, J. S. (2004). Temporal pattern of Africanization in a feral honey bee population from Texas inferred from mitochondrial DNA. *Evolution*, *58*(5), 1047-1055.

Pinto, M. A., Muñoz, I., Chávez-Galarza, J., & De la Rúa, P. (2012). The Atlantic side of the Iberian Peninsula: a hot-spot of novel African honey bee maternal diversity. *Apidologie*, *43*(6), 663-673.

Pinto, M. A., Henriques, D., Neto, M., Guedes, H., Muñoz, I., Azevedo, J. C., & De la Rúa, P. (2013). Maternal diversity patterns of Ibero-Atlantic populations reveal further complexity of Iberian honey bees. *Apidologie*, *44*(4), 430-439.

Pinto, M. A., Henriques, D., Chávez-Galarza, J., Kryger, P., Garnery, L., van der Zee, R., Dahle, B., Soland-Reckweg, G., De la Rúa, P., Dall'Olio, R., Carreck, N. L. & Johnston, J. S. (2014). Genetic integrity of the Dark European honey bee (*Apis mellifera mellifera*) from protected populations: a genome-wide assessment using SNPs and mtDNA sequence data. *Journal of Apicultural Research*, *53*(2), 269-278.

Pirk, C. W. W., Neumann, P., & Ratnieks, F. L. W. (2003). Cape honeybees, *Apis mellifera capensis*, police worker-laid eggs despite the absence of relatedness benefits. *Behavioral Ecology*, *14*(3), 347-352.

Pitchers, W., Pool, J. E., & Dworkin, I. (2013). Altitudinal clinal variation in wing size and shape in African *Drosophila melanogaster*: one cline or many?. *Evolution*, *67*(2), 438-452.

Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). Global pollinator declines: trends, impacts and drivers. *Trends in ecology & evolution*, *25*(6), 345-353.

Queiroz, G., Pacheco, J. M., Gaspar, J. L., Aspinall, W. P., Guest, J. E., & Ferreira, T. (2008). The last 5000 years of activity at Sete Cidades volcano (São Miguel Island, Azores): Implications for hazard assessment. *Journal of Volcanology and Geothermal Research*, *178*(3), 562-573.

Raffiudin, R., & Crozier, R. H. (2007). Phylogenetic analysis of honey bee behavioral evolution. *Molecular phylogenetics and evolution*, *43*(2), 543-552.

Rasmont, P., Scholl, A., De Jonghe, R., Obrecht, E., & Adamski, A. (1986). Identité et variabilité des mâles de bourdons du genre *Bombus Latreille* sensu stricto en Europe occidentale et centrale (Hymenoptera, Apidae, Bombinae). *Revue suisse de Zoologie*, *93*(3), 661-682.

Ricklefs, R. E., & Miles, D. B. (1994). Ecological and evolutionary inferences from morphology: an ecological perspective. *Ecological morphology: integrative organismal biology. University of Chicago Press, Chicago*, *101*, 13-41.

Rohlf, F. J., & Marcus, L. F. (1993). A revolution morphometrics. *Trends in Ecology & Evolution*, *8*(4), 129-132.

Rohlf, F. J. (2002). Geometric morphometrics and phylogeny. *Morphology, shape and phylogeny*, 175-193.

Rohlf F.J. (2005). tpsRelw, version 2.04. Department of Ecology and Evolution, State University of New York, Stony Brook.

Rohlf, F.J., (2013). TPSdig, v2.17. State University at Stony Brook, NY.

Rortais, A., Arnold, G., Alburaki, M., Legout, H., & Garnery, L. (2011). Review of the DraICOI-COII test for the conservation of the black honeybee (*Apis mellifera mellifera*). *Conservation Genetics Resources*, *3*(2), 383-391.

Rosenkranz, P., Aumeier, P., & Ziegelmann, B. (2010). Biology and control of *Varroa destructor*. *Journal of invertebrate pathology*, *103*, 96-119.

Ruttner, F. (1975). Races of bees. *The hive and the honey bee*, *740*, 19-38.

Ruttner, F. (1988). Biogeography and taxonomy of honey bees Springer. *Berlin, Germany*.

Ruttner, F. (1992) Naturgeschichte der honigbienen. Ehrenwirth, München, 357 pp.

Santos, R. S., Hawkins, S., Monteiro, L. R., Alves, M., & Isidro, E. J. (1995). Marine research, resources and conservation in the Azores. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *5*(4), 311-354.

Savriama, Y., Gómez, J. M., Perfectti, F., & Klingenberg, C. P. (2012). Geometric morphometrics of corolla shape: dissecting components of symmetric and asymmetric variation in *Erysimum mediohispanicum* (Brassicaceae). *New Phytologist*, *196*(3), 945-954.

Secretaria Regional da Agricultura e Ambiente (SRAA) (2016). Situação Epidemiológica. Região Autónoma dos Açores.

Segura, S. K. (2016). Estudio de los patógenos de *Apis Mellifera* Linnaeus, 1958 en el Archipiélago de Azores y la isla de Madeira (Master thesis (Msc)), Zoologia, Universidad Complutense de Madrid, Madrid, Spain (40 pp).

Sheppard, W. S., Arias, M. C., Grech, A., & Meixner, M. D. (1997). *Apis mellifera ruttneri*, a new honey bee subspecies from Malta. *Apidologie*, *28*, 287-294.

Sheppard, W., & Meixner, M. (2003). *Apis mellifera pomonella*, a new honey bee subspecies from Central Asia. *Apidologie*, *34*(4), 367-375.

Shipunov, A. B., & Bateman, R. M. (2005). Geometric morphometrics as a tool for understanding *Dactylorhiza* (Orchidaceae) diversity in European Russia. *Biological Journal of the Linnean Society*, *85*(1), 1-12.

Slice, D. E. (2007). Geometric morphometrics. *Annual Review of Anthropology*, *36*, 261-281.

Smith, U. E., & Hendricks, J. R. (2013). Geometric morphometric character suites as phylogenetic data: extracting phylogenetic signal from gastropod shells. *Systematic Biology*, *62*, 366–385.

Stevanovic, J., Stanimirovic, Z., Radakovic, M., & Kovacevic, S. R. (2010). Biogeographic study of the honey bee (*Apis mellifera* L.) from Serbia, Bosnia and Herzegovina and Republic of Macedonia based on mitochondrial DNA analyses. *Russian Journal of Genetics*, *46*(5), 603-609.

Sušnik, S., Kozmus, P., Poklukar, J., & Meglic, V. (2004). Molecular characterisation of indigenous *Apis mellifera carnica* in Slovenia. *Apidologie*, *35*(6), 623-636.

Szalanski, A. L., Tripodi, A. D., Trammel, C. E., & Downey, D. (2016). Mitochondrial DNA genetic diversity of honey bees, *Apis mellifera*, in Hawaii. *Apidologie*, *47*(5), 679-687.

Tamura K, Stecher G, Peterson D et al. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, 30, 2725–2729.

Techer, M. A., Clémencet, J., Simiand, C., Portlouis, G., Reynaud, B., & Delatte, H. (2016). Genetic diversity of the honey bee (*Apis mellifera* L.) populations in the Seychelles archipelago. *Insect Conservation and Diversity*, *9*(1), 13-26.

Techer, M. A., Clémencet, J., Simiand, C., Preeaduth, S., Azali, H. A., Reynaud, B., & Hélène, D. (2017). Large-scale mitochondrial DNA analysis of native honey bee *Apis mellifera* populations reveals a new African subgroup private to the South West Indian Ocean islands. *BMC Genetics*, *18*(1), 53.

Techer, M. A., Clémencet, J., Turpin, P., Volbert, N., Reynaud, B., & Delatte, H. (2015). Genetic characterization of the honey bee (*Apis mellifera*. *Apidologie*, *46*(4), 445-454.

Tofilski, A. (2008). Using geometric morphometrics and standard morphometry to discriminate three honey bee subspecies. *Apidologie*, *39*(5), 558-563.

Villemant, C., Simbolotti, G., & Kenis, M. (2007). Discrimination of *Eubazus* (Hymenoptera, Braconidae) sibling species using geometric morphometrics analysis of wing venation. *Systematic Entomology*, *32*(4), 625-634.

Viscosi, V., & Cardini, A. (2011). Leaf morphology, taxonomy and geometric morphometrics: a simplified protocol for beginners. *PloS one*, *6*(10), e25630.

Wallberg, A., Han, F., Wellhagen, G., Dahle, B., Kawata, M., Haddad, N., Simões, Z. L. P., Allsopp, M.H., Kandemir, I., De la Rúa, P., Pirk, C. W. & Webster, M.T. (2014). A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honey bee Apis mellifera. *Nature genetics*, *46*(10), 1081-1088.

Wappler, T., De Meulemeester, T., Murat Aytekin, A., Michez, D., & Engel, M. S. (2012). Geometric morphometric analysis of a new Miocene bumble bee from the Randeck Maar of southwestern Germany (Hymenoptera: Apidae). *Systematic Entomology*, *37*(4), 784-792.

Whitfield, C. W., Behura, S. K., Berlocher, S. H., Clark, A. G., Johnston, J. S., Sheppard, W. S., Smith, D. R., Suarez, A. V., Weaver, D. & Tsutsui, N. D. (2006). Thrice out of Africa: ancient and recent expansions of the honey bee, *Apis mellifera*. *Science*, *314*(5799), 642-645.

Zelditch, M., Swiderski, D.L., Sheets, H.D. & Fink, W.L. (2004). *Geometric Morphometrics for Biologists: A Primer*. Elsevier Academic Press, London.

7. Supplementary Material

Table S1. Distribution of the 31 susbspecies of *A. mellifera* according the respective evolutionary lineage: A, M, C and O.

Lineage	Subspecie	Distribution
M	*Apis mellifera mellifera* Linnaeus (1958)	France to Scandinavia and from British Isles to Central Europe, including Central Russia
	Apis mellifera iberiensis Engle (1999)	Iberian peninsula, Balearic islands and Macaronesia Islands
	Apis mellifera sinisxinyuan Chen et al. (2016)	Yili River Valley in Xinjiang Uygur Autonomous Region, western China
A	*Apis mellifera adansonii* Latreille (1804)	Western Africa ranging from Niger in the north, east to Senegal and as far south as Democratic Republic of Congo
	Apis mellifera siciliana Grassi (1880)	Island of Sicily
	Apis mellifera intermissa Maa (1953)	Northen coast of Africa, Marroco and Tunisia
	Apis mellifera jemenitica Ruttner (1988)	Chad, Oman Saudi Arabia, Somalia, Sudan and Yemen
	Apis mellifera lamarckii Cockerell (1906)	Narrow range across the Egyptian Nile Valey
	Apis mellifera sahariensis Baldensperger (1923)	Southern side of the Atlas range in oases from Marocco to Algeria
	Apis mellifera scutellata Lapeletier (1836)	South Africa northward along the eastern half of the continent to about Ethiopia
	Apis mellifera capensis Escholtz (1822)	Cape Region in South Africa
	Apis mellifera montícola (Smith (1961)	Mountains of Eastern Africa (Kenya and Tanzania)
	Apis mellifera unicolor Latreille (1804)	Island of Madagascar
	Apis mellifera litorea Smith (1961)	Eastern coast of tropical Africa from Southern Kenya to Mozambique
	Apis mellifera ruttneri Sheppard et al. (1977)	Island of Malta
	Apis mellifera simensis Meixner et al. (2011)	Mountain systems of Ethiopia
C	*Apis mellifera ligústica* Spinola (1806)	Italian Peninsula, confined by the Alps in the north and Mediterranean Sea in south
	Apis mellifera carnica Pollmann (1879)	Austria, Slovenia, Croatia, Bosnia-Herzegovina, Albania, Serbia, Hungary and Romania
	Apis mellifera macedonica Ruttner (1988)	Bugaria, the Former Yugoslav Republic of Macedonia (FYROM), Greece, Romania, Ukraine and Turkey
	Apis mellifera cecropia Kiesenwetter (1860)	Southern Greece including Peloponnese and Surrounding Aegean islands
O	*Apis mellifera syriaca Skorikov (1929)*	Eastern shores of the Mediterranean Sea; North from Syria to the Negev Desert in the South /Syria, Lebanon, Jordan and Israel
	Apis mellifera remipes Gerstäcker (1862)	Armenia
	Apis mellifera meda Skorikov (1929)	Iran, Iraq, Southern Turkey and Northen Syria
	Apis mellifera anatoliaca Maa (1953)	Turkey

	Apis mellifera caucasia Pollmann (1889)	Caucasus Mountains
	Apis mellifera cypria Pollmann (1879)	Island of Cyprus
	Apis mellifera adami Ruttner (1975)	Island of Crete
	Apis mellifera pomonella Sheppard and Meixner (2003)	Tie Shan Mountains in Central Asia along east-west orientation from South central Kasakhstan to western China
?	*Apis mellifera sossimai Engle (1999)*	Ukraine with the exception of Crimea and Northern regions of the Caucasus Mountains
	Apis mellifera taurica Apaltov (1938)	Along the north-central shores of the Black Sea; in the Crimea.
	Apis mellifera artemisia Engle (1999)	Central Russian Steppes

Table S2. PCR-RFLP assay that have been used to identify honey bee matrilineal origins (Adapted from Meixner et al. 2013).

Genes	Restriction Enzyme	Authors
Cytochrome b	*Bgl*II	Crozier et al. 1991
Ls rRNA	*Eco*RI	Hall and Smith, 1991
COI	*Hinc*II	Hall and Smith, 1991
	*Xba*I	Hall and Smith, 1991
	*Hinf*I	Nielsen et al. 2000
	*Nco*I	Bouga et al. 2005
		Stevanovic et al. 2010
	*Sty*I	Bouga et al. 2005
		Stevanovic et al. 2010
	*Ssp*I	Ivanova et al. 2010
ND5	*Alu*I	Bouga et al. 2005
	*Hinc*II	Ivanova et al. 2010
	*Fok*I	Ivanova et al. 2010
COI-COII	D*ra*I	Garnery et al. 1993
		Garnery et al. 1995
		Garnery et al. 1998
		Franck et al. 1998
		Franck et al. 2001
		De La Rúa et al. 1998
		De La Rúa et al. 2001a
		De La Rúa et al. 2001b
		De La Rúa et al. 2002
		De La Rúa et al. 2004
		De La Rúa et al. 2005
		Susnik et al. 2004
		Jensen et al. 2005
		Miguel et al. 2007
		Cánovas et al. 2008
		Rortais et al. 2011
		Pinto et al. 2012
		Pinto et al. 2013

Table S3. Beekeeping statistics provided by "Secretaria Regional da Agricultura e do Ambiente, Região Autónoma dos Açores" for 2014 and 2015, the years when sampling was carried out.

Year	Island	Number of Beekeepers	Number of Apiaries	Number of Colonies
2014	Santa Maria	37	44	250
	São Miguel	133	313	2035
	Terceira	71	111	972
	Graciosa	7	11	195
	São Jorge	20	25	176
	Pico	48	79	907
	Faial	30	48	356
	Flores	19	20	120
	Total	**365**	**651**	**5011**
2015	Santa Maria	40	48	319
	São Miguel	194	322	2.345
	Terceira	75	130	1.172
	Graciosa	7	12	188
	São Jorge	20	31	179
	Pico	52	85	943
	Faial	32	57	471
	Flores and Corvo	17	18	127
	TOTAL	**437**	**703**	**5.744**

Table S4. Procrustes (below the diagonal) and Mahalanobis (above the diagonal) distances between populations, obtained with the DFA data for the wings of honey bee from the eight islands of the Azores Archipelago. All distances were highly significant (*P-value <0.0001).*

Mahalanobis / Procrustes	Faial	Flores	Graciosa	Pico	São Jorge	Santa Maria	São Miguel	Terceira
Faial	-	3.7856	5.9052	2.1520	3.9004	5.8700	4.4598	4.1336
Flores	0.0082	-	10.700	4.0289	6.1485	5.9524	4.8833	6.6033
Graciosa	0.0153	0.0164	-	6.8601	11.7358	9.5393	7.9601	6.2250
Pico	0.0064	0.0090	0.0158	-	3.4654	6.2696	4.5455	4.5086
São Jorge	0.0084	0.0082	0.0183	0.0066	-	6.8082	5.7911	4.8198
Santa Maria	0.0146	0.0142	0.0234	0.0149	0.0152	-	4.1287	6.1570
São Miguel	0.0104	0.0126	0.0173	0.0117	0.0140	0.0092	-	5.8516
Terceira	0.0091	0.0098	0.0126	0.0087	0.0103	0.0155	0.0121	-

Table S5. Procrustes (below the diagonal) and Mahalanobis (above the diagonal) distances, obtained with the DFA data for the wings of honey bee from the eight islands of the Azores Archipelago and the reference samples *A. m. iberiensis*, *A. m. ligustica* and *A. m. carnica*. All distances were highly significant (*P-value <0.0001)*.

Mahalanobis / Procrustes	Faial	Flores	Graciosa	Pico	S.Jorge	S. Maria	S.Miguel	Terceira	iberiensis	ligustica	carnica
Faial	-	3.7813	5.9051	2.1520	3.9004	5.8702	4.4634	4.1335	5.4364	10.0291	10.6700
Flores	0.0082	-	10.6850	4.0316	6.1041	5.8913	4.9287	6.6076	6.2066	15.3032	16.8754
Graciosa	0.0153	0.0164	-	6.8600	11.7353	9.5388	7.9125	6.2249	7.9574	11.5917	15.1662
Pico	0.0064	0.0090	0.0158	-	3.4653	6.2697	4.5332	4.5086	5.1065	12.4560	11.9548
S.Jorge	0.0084	0.0082	0.0183	0.0066	-	6.8083	5.7959	4.8199	5.2651	18.8785	16.3391
S.Maria	0.0146	0.0143	0.0234	0.0149	0.0152	-	4.1308	6.1571	5.5985	13.4600	16.4584
S.Miguel	0.0104	0.0127	0.0173	0.0117	0.0140	0.0092	-	5.8420	4.3707	12.0798	13.3384
Terceira	0.0091	0.0098	0.0126	0.0087	0.0103	0.0155	0.0121	-	6.5071	10.8014	13.9796
Iberiensis	0.0133	0.0142	0.0203	0.0126	0.0147	0.0107	0.0101	0.0136	-	10.3143	11.3880
ligustica	0.0217	0.0228	0.0133	0.0218	0.0231	0.0310	0.0256	0.0192	0.0272	-	6.5054
carnica	0.0244	0.0261	0.0182	0.0234	0.0260	0.0322	0.0281	0.0219	0.0281	0.0110	-

Table S6. Procrustes (below the diagonal) and Mahalanobis (above the diagonal) distances, generated with the DFA data obtained for populations from the Azores archipelago, continental Portugal, and Spain. All distances were highly significant (*P-value <0.0001)*.

Mahalanobis / Procrustes	Spain	Portugal	Azores
Spain	-	1.7206	4.3966
Portugal	0.0054	-	3.6360
Azores	0.0110	0.0096	-

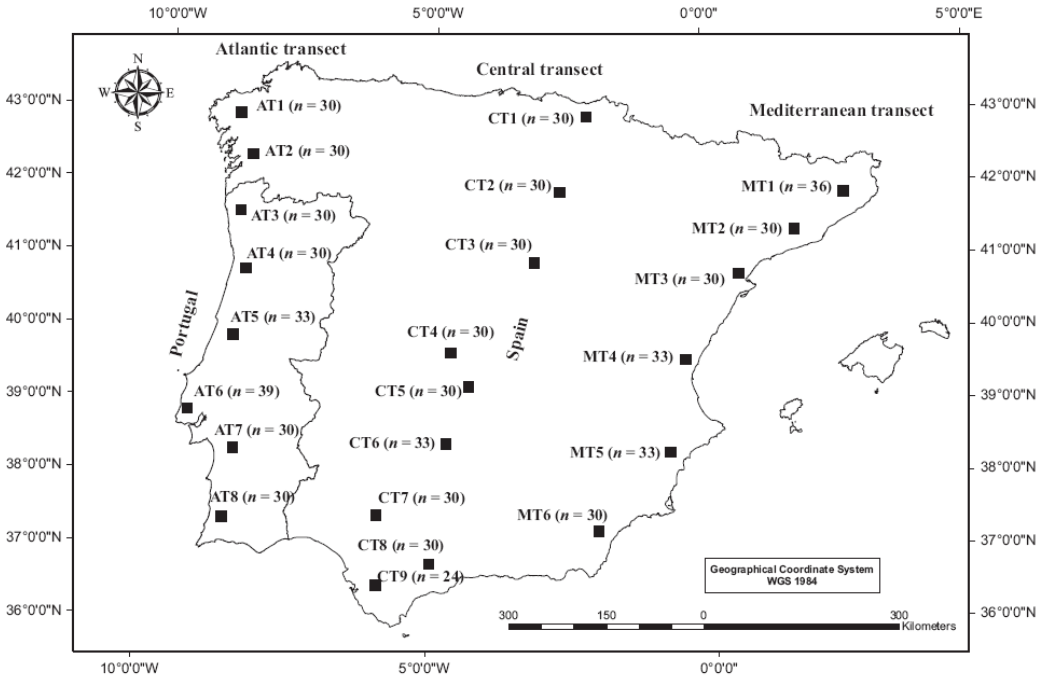

Figure S1. Map of the Iberian Peninsula w ith the location of the 237 apiaries sampled across three transects: Atlantic transect (AT), Central transect (CT), and Mediterranean transect (MT). Map obtained from Chávez-Galarza et al. 2013.

tRNAleu
P element
d1
P1
Haplotypes
Sub-Lineage
Lineage
C1
C2j
C3b
A64'
A30
A30d
C
AIII
A
Q1 element
Q2 element
Q3 element
Cox2

Figure S2. Sequence alignment of the intergenic $tRNA^{leu}$-cox2 region showing variation among haplotypes of C and A lineages. Substitution sites are numbered from 1 to 10. Indels (showed with a dash) are indicated by letters from "a" to "f" with a grey background. *Dra*I recognition sites (TTTAAA) are indicated by nucleotides in black and white background. Unidentified nucleotides are represented by a bold N with a grey background. The absence of the P element characterizes lineage C. Sub-lineage AIII is identified by the P1 element (denoted by the "*d1"* deletion marked in grey). The haplotype A30, with just a recognition site, was identified not only with the sequence but also with PCR-RFLP band pattern in a 4% wide-range agarose gel.